

(19) 日本国特許庁 (J P)

(12) 特 許 公 報 (B 2)

(11) 特許番号

第2983194号

(45) 発行日 平成11年(1999)11月29日

(24) 登録日 平成11年(1999)9月24日

(51) Int.Cl. <sup>8</sup>	識別記号	P I	
G 1 1 B 20/10		G 1 1 B 20/10	B
	3 2 1		3 2 1 Z
G 0 9 B 5/04		G 0 9 B 5/04	
G 1 1 B 20/12	1 0 2	G 1 1 B 20/12	1 0 2

請求項の数6 (全 20 頁)

(21) 出願番号 特願平9-294334

(22) 出願日 平成9年(1997)10月27日

(65) 公開番号 特開平11-134791

(43) 公開日 平成11年(1999)5月21日

審査請求日 平成9年(1997)10月27日

(73) 特許権者 000104179  
カナース・データー株式会社  
東京都千代田区京神田1丁目10番7号

(72) 発明者 関口 博司  
東京都千代田区京神田1-10-7 カナース・データー株式会社内

(74) 代理人 弁理士 長谷川 芳樹 (外3名)

審査官 小松 正

(58) 調査した分野(Int.Cl.<sup>8</sup>, DB名)  
G11B 20/10  
G09B 5/04

(54) 【発明の名称】 音声記録媒体、その再生方法、及び音声再生中の割込み処理方法

1

(57) 【特許請求の範囲】

【請求項1】 所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された音声情報列が記録された第1領域と、該音声情報列における各区画ごとに、該各区画の前半部分と後半部分との境界を示す境界位置識別情報が記録された第2領域とを、少なくとも備えた音声記録媒体の音声情報再生中に発生した、種々の再生方法を指示する割込み要求の処理を実行するための音声再生中の割込み処理方法において、前記区画のうち所定区画の再生中に発生した割込み要求に対し、該割込み要求が、該再生中であった区画の先頭位置から境界位置までの前半部分で発生したか、あるいは該再生中であった区画の境界位置から次区画の先頭位

2

置までの後半部分で発生したかを判断し、前記割込み要求の発生タイミングに関する判断結果に基づいて、前記再生中であった区画の前半部分で発生したと判断された前記割込み要求が対象とする区画として、該再生中であった区画の前区画を特定するか、あるいは前記再生中であった区画の後半部分で発生したと判断された前記割込み要求が対象とする区画として、前記再生中であった区画自身を特定することを特徴とする音声再生中の割込み処理方法。

【請求項2】 所定の記録媒体に記録された音声情報を再生する、音声記録媒体の再生方法において、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された第1音声情報列が記録さ

(2)

特許2983194

3

れた第1領域と、該第1音声情報列と等価でかつ遅延した発音で構成された音声情報列であって、該第1音声情報列の各区画の音声情報と等価な音声情報ごとに可変長の区画に分割された第2音声情報列が記録された第2領域と、該第1及び第2音声情報列を該音声再生手段で切換えて再生すべく、切換え可能な各区画を当該記録媒体における該各区画の記録位置で示す記録位置識別情報が記録された第3領域と、そして、該第1及び第2音声情報列における各区画ごとに、該各区画の前半部分と後半部分との境界を示す境界位置識別情報が記録された第4領域とを、少なくとも備えた音声記録媒体を用意し、前記区画のうち所定区画の再生中に発生した前記第1及び第2音声情報列間の切換え再生に関する第1の割込み要求に対し、該第1の割込み要求が、該再生中であった区画の先頭位置から境界位置までの前半部分で発生したか、あるいは該再生中であった区画の境界位置から次区画の先頭位置までの後半部分で発生したかを判断し、前記第1の割込み要求の発生タイミングに関する判断結果に基づいて、前記再生中であった区画の前半部分で発生したと判断された該第1の割込み要求が対象とする区画として、該再生中であった区画の前区画を特定するか、あるいは前記再生中であった区画の後半部分で発生したと判断された前記第1の割込み要求が対象とする区画として、前記再生中であった区画自身を特定し、前記第1の割込み要求が対象とする区画として特定された区画に対応する区画であって、前記第1及び第2音声情報列のうち、該第1の割込み要求によって指示された音声情報列の区画の先頭から、音声再生を再開することを特徴とする、音声記録媒体の再生方法。

【請求項3】 前記区画のうち所定区画の再生中に発生したリビート再生に関する第2の割込み要求に対し、該第2の割込み要求が、該再生中であった区画の先頭位置から境界位置までの前半部分で発生したか、あるいは該再生中であった区画の境界位置から次区画の先頭位置までの後半部分で発生したかを判断し、前記第2の割込み要求の発生タイミングに関する判断結果に基づいて、前記再生中であった区画の前半部分で発生したと判断された該第2割込み要求が対象とする区画として、前記再生中であった区画の前区画を特定するか、あるいは前記再生中であった区画の後半部分で発生したと判断された前記第2の割込み要求が対象とする区画として、前記再生中であった区画自身を特定し、前記第1及び第2音声情報列のうち、再生中であった区画を含む音声情報列であって、前記第2の割込み要求が対象とする区画として特定された区画の先頭から、音声再生を再開することを特徴とする、請求項2記載の音声記録媒体の再生方法。

【請求項4】 所定の記録媒体に記録された音声情報を再生する、音声記録媒体の再生方法において、所定の音声再生手段で再生出力されるべき複数の単語列

4

から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された音声情報列が記録された第1領域と、該音声情報列における各区画ごとに、該各区画の前半部分と後半部分との境界を示す境界位置識別情報が記録された第2領域とを、少なくとも備えた音声記録媒体を用意し、

前記区画のうち所定区画の再生中に発生したリビート再生に関する割込み要求に対し、該割込み要求が、該再生中であった区画の先頭位置から境界位置までの前半部分で発生したか、あるいは該再生中であった区画の境界位置から次区画の先頭位置までの後半部分で発生したかを判断し、

前記割込み要求の発生タイミングに関する判断結果に基づいて、前記再生中であった区画の前半部分で発生したと判断された前記割込み要求が対象とする区画として、該再生中であった区画の前区画を特定するか、あるいは前記再生中であった区画の後半部分で発生したと判断された前記割込み要求が対象とする区画として、前記再生中であった区画自身を特定し、

前記割込み要求が対象とする区画として特定された区画の先頭から、音声再生を再開することを特徴とする、音声記録媒体の再生方法。

【請求項5】 所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された音声情報列が記録された第1領域と、

前記音声情報列における各区画ごとに、該各区画の前半部分と後半部分との境界を示す境界位置識別情報が記録された第2領域とを、少なくとも記録された音声記録媒体、

【請求項6】 所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された第1音声情報列が記録された第1領域と、

前記第1音声情報列と等価でかつ遅延した発音で構成された音声情報列であって、該第1音声情報列の各区画の音声情報と等価な音声情報ごとに可変長の区画に分割された第2音声情報列が記録された第2領域と、

前記第1及び第2音声情報列を該音声再生手段で切換えて再生すべく、切換え可能な各区画を当該記録媒体における該各区画の記録位置で示す記録位置識別情報が記録された第3領域と、そして、

前記第1及び第2音声情報列における各区画ごとに、該各区画の前半部分と後半部分との境界を示す境界位置識別情報が記録された第4領域とを、少なくとも備えた音声記録媒体、

【発明の詳細な説明】

(3)

特許2983194

5

【0001】

【発明の属する技術分野】この発明は、CD-ROM、DAT等の媒体に少なくとも音声情報を含む種々の情報が記録された音声記録媒体と、このような音声記録媒体にあらかじめ記録された音声情報を再生する再生方法と、音声再生中に発生する、種々の再生方法を指示する割込み要求を処理するための割込み処理方法に関するものである。

【0002】

【従来の技術】従来から、英会話等の語学の独習用、詩吟の練習用、法律の独習用、その他の目的のために、カセットテープ等の記録媒体に音声情報が記録された教材が種々提供されている。ここで、英会話の独習用の教材を例に説明すると、従来の主な記録媒体は、例えば一連の英語の発声（音声情報）が記録されたカセットテープ（又はレコード）であり、学習者はこのテープ教材とテキストとを組み合わせて使用していた。なお、このような教材には、初級用から上級用まで種々のレベルが用意されている。

【0003】また、日本国特許第2581700号には、複数の区画に区分された上級者学習用に適した音声情報列が記録された第1領域と、これら各区画に対応した等価な区画からなる初級者学習用に適した音声情報列が記録された第2領域と、該上級者学習用及び初級者学習用の各音声情報列の対応する各区画の関係を、これら音声情報列の各区画の記録媒体における記録位置で示す情報が記録された第3領域とを、少なくとも備えたCD-ROM等の情報記録媒体、及びこのような構造を備えた情報記録媒体の切替え再生等を含む再生方法が提案されている。

【0004】一方、特開平5-224581号公報には、語学学習用CDを利用した音声再生中のリピート機能を実現するための技術が提案されている。すなわち、この技術は通常のヒヤリング学習中における学習者からの指示遅れを感知する遅延感知機能に関するものであり、具体的には、あらかじめインデックスナンバーが付された1センテンス（文）ごとの音声情報列が順次記録された媒体を所定の再生装置を用いて再生している。例えば、このような再生装置にリピート再生ができる機能が設けられていた場合であっても、難解なセンテンスや、聞き取りにくい発音を含んだセンテンスを聞いた直後等では、リピート指示（学習者による機器操作）が遅れて次のセンテンスの再生が始まってしまう、学習者は既に聞き終わったセンテンスを聞くことができなくなるという不具合があった。そこで、この従来技術では、図19に示されたように、順次A、B、C、Dの各時点で再生対象となるセグメントのインデックスナンバーが順次切替えられることにより、再生対象となるセグメントの音声情報が再生されるが、同時に次の音声情報の再生開始から時間 $\Delta$ が経過まで前のセグメントを示すインデ

6

ックスナンバー（リピート再生する音声情報を示すインデックスナンバー）も所定の保持手段により保持される。換言すれば、この従来技術は再生対象であるセグメントのインデックスナンバーとリピート再生の対象であるべきセグメントのインデックスナンバーとを独立に保持し、これら各インデックスナンバーの切替えタイミングを所定時間 $\Delta$ だけずらすよう時間管理している。このように構成することにより、学習者からのリピート要求が発行された時点での保持されていたインデックスナンバーに対応するジャンプ信号（再生開始位置を指示する情報）が発生させ、学習者からのリピート要求の遅れに起因した問題（指示遅れ）を回避している。

【0005】

【発明が解決しようとする課題】以上のように、従来の音声再生方法では、リピート機能のような操作者（例えば、語学学習者）からの指示遅れをあらかじめ考慮し、操作者が意図したセンテンスからリピート再生が行えるよう、該操作者からの指示の遅延感知が行われている。

【0006】しかしながら、このような遅延感知機能は所定のセンテンスが順次規則的に記録された場合を想定した技術である。したがって、例えば映画における出演者の会話、日常の生活環境における会話等のように、長さの異なる複数のセンテンスから構成され、また、各センテンス（各会話者の音声情報）の間に、音声再生されていない状況、雑音のみが再生されている状況、音楽（BGM）のみが再生されている状況等のランダムに発生する無音声期間が長時間（例えば1秒以上）存在し得る一連の音声情報の再生では、時間管理による遅延感知機能では操作者である学習者が希望する動作を的確に実現するのは困難である。すなわち、このような状況下では操作者から意図したリピート対象センテンスと、実際にリピート再生されるセンテンスとが異なる場合が起こり、結果的に語学学習等では機器操作に学習者の注意が奪われ十分な学習効果が低下してしまう（集中力が散漫になる）という課題があった。

【0007】そこで、この発明は上述のような課題を解決するためになされたものであり、長さの異なる複数のセンテンスから構成された一連の音声情報を再生する場合であっても、的確に操作者が意図する音声情報の再生を可能にする構造を備えた音声記録媒体、その再生方法、及び音声再生中の割込み処理方法を提供することを目的としている。

【0008】

【課題を解決するための手段】この発明は、映画における出演者の会話、日常の生活環境における会話等のように、長さの異なる複数のセンテンスから構成され、また、各センテンス（各会話者の音声情報）の間に、音声再生されていない状況、雑音のみが再生されている状況、音楽（BGM）のみが再生されている状況等のランダムに発生する無音声期間が存在し得る一連の音声情報

(4)

特許2983194

7

の再生においても、操作者が意図する音声情報の再生を、指示遅れも考慮して的確に実現する技術に関するものである。

【0009】特に、再生対象である音声記録媒体には、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された1種類又は2種類以上の音声情報列が少なくとも記録されている。なお、この明細書において、音声情報とは、実際に聞き取れる音（音声）を含む概念及び情報であり、音声情報列とは、実際にCD-ROM等の記録媒体に記録されたデータ自体を意味する。

【0010】具体的に、この発明に係る音声記録媒体（音声記録媒体に係る第1実施態様）は、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目（センテンスの終りやセンテンス中の一息つける、発生上あるいは文法上の区切り）でそれぞれ分割された音声情報ごとに可変長の区画に区分された音声情報列が記録された第1領域と、該音声情報列の各区画ごとに、該各区画の前半部分と後半部分との境界を示す境界位置識別情報が記録された第2領域とを、少なくとも備えている。

【0011】なお、当該音声情報記録媒体は、複数種類の音声情報列が記録された複数の領域を備えてもよい。すなわち、音声再生中の遅延感知機能は、上述の日本国特許第2581700号に示されたように複数種類の音声情報の切替え再生可能な場合、この発明に係る音声記録媒体（音声記録媒体に係る第2実施態様）は、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された第1音声情報列が記録された第1領域と、及び該第1音声情報列と等価でかつ遅延した発音で構成された音声情報列であって、該第1音声情報列の各区画の音声情報と等価な音声情報ごとに可変長の区画に区分された第2音声情報列が記録された第2領域と、該第1及び第2音声情報列を該音声再生手段で切替えて再生すべく、切替え可能な各区画を当該記録媒体における該各区画の記録位置で示す記録位置識別情報が記録された第3領域と、そして、該1及び第2音声情報列における各区画ごとに、該各区画の前半部分と後半部分の境界を示す境界位置識別情報が記録された第4領域とを、少なくとも備えている。

【0012】また、この発明に係る音声記録媒体は、上述のいずれの構成であっても、区画に関する情報として、各区画中の発音状態に応じて（発音のまとまりごとに）区切られた前半部分と後半部分の境界を示す境界位置識別情報が、当該音声記録媒体の所定領域に記録され

8

たことを特徴としている。

【0013】一方、この発明に係る音声再生方法は、音声再生中に発生する種々の再生方法を指示する割込み要求に対し、各区画にあらかじめ設定された上記境界位置識別情報を利用して、該割込み要求の対象を特定している。なお、割込み要求の再生指示には、リピート機能、上記音声記録媒体の各領域に格納された複数の音声情報を切替え再生する切替え再生機能、あるいはこれらの組み合わせ等が少なくとも含まれる。

【0014】上述のリピート機能を実現する場合、この発明に係る音声記録媒体の再生方法（再生方法に係る第1の実施態様）では、まず、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された音声情報列が記録された第1領域と、該音声情報列の各区画ごとに、発音状態によって区切られた該区画の前半部分と後半部分の境界を示す境界位置識別情報が記録された第2領域とを、少なくとも備えた音声記録媒体が用意される。そして、記録された複数の区画のうち所定区画の再生中に操作者からのリピート再生に関する割込み要求が発行されると、この割込み要求の発生タイミングが判断される。具体的には、再生中であった区画の先頭位置から境界位置までの該区画の前半部分で割込み要求が発生したか、あるいは該再生中であった区画の境界位置から次区画の先頭位置までの該区画の後半部分で割込みが発生したかが判断される。ここで、上記割込み要求が再生中であった区画の前半部分で発生したと判断されると、割込み要求が対象とする区画として再生中であった区画の前区画が特定される。一方、上記割込み要求が再生中であった区画の後半部分で発生したと判断されると、割込み要求が対象とする区画として再生中であった区画自身が特定される。

【0015】再生方法に係る第1実施態様では、以上のような割込み処理により特定された区画（操作者が意図した対象区画として特定された区画）の先頭から、音声再生を再開することによりリピート機能を実現する。

【0016】一方、当該記録媒体に記録された複数の音声情報列を切替え再生する場合、この発明に係る音声記録媒体の再生方法（再生方法に係る第2の実施態様）では、まず、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区画に区分された第1音声情報列が記録された第1領域と、該第1音声情報列と等価でかつ遅延した発音で構成された音声情報列であって、該第1音声情報列の各区画の音声情報と等価な音声情報ごとに可変長の区画に分割された第2音声情報列が記録された第2領域と、該第1及び第2音声情報列を該音声再生手段で切替えて再生すべく、切替え可能な各区画を当

(5)

特許2983194

9

該記録媒体における該各区画の記録位置で示す記録位置識別情報が記録された第3領域と、そして、該第1及び第2音声情報列における各区分ごとに、発音状態によって区切られた該各区分の前半部分と後半部分の境界を示す境界位置識別情報が記録された第4領域とを、少なくとも備えた音声記録媒体が用意される。そして、これらの区分のうち所定区分の再生中に第1及び第2音声情報列間での切換え再生に関する割込み要求が発生した場合、一旦割込み要求の発生タイミングが判断される。すなわち、この割込み要求が、再生中であった区分の先頭位置から境界位置までの該区分の前半部分で発生したか、あるいは該再生中であった区分の境界位置から次区分の先頭位置までの該区分の後半部分で発生したかが判断され、この判断結果に基づいて割込み対象となる区分が特定される。特定される区分は、この割込み要求が再生中であった区分の前半部分で発生したと判断された場合、該区分の前区分が特定され、割込み要求が再生中であった区分の後半部分で発生したと判断された場合には、該区分自身が特定される。

【0017】再生方法に係る第2の実施態様では、以上のような割込み処理により操作者が意図した指示対象としての区分が特定されると、割込み要求が対象とする区分として特定された区分に対応する区分であって、第1及び第2音声情報列のうち上記割込み要求によって指示された音声情報列の区分の先頭から、音声再生が再開されることにより切替え再生機能を実現する。

【0018】なお、以上のように切替え再生可能な音声記録媒体の再生方法であっても、上述のリピート再生等の種々の機能をさらに備えることは可能である。

【0019】また、この発明に係る音声記録媒体は、ポータブルCDプレーヤーやCD-ROMドライブ等を備えたパーソナルコンピュータにより再生可能な情報記録媒体を含む。したがって、このような種々のデバイスにより再生されるべく、各音声情報の再生タイミングに対応して再生可能な画像情報、文字情報等を含んでもよい。

【0020】

【発明の実施の形態】以下、この発明の一実施例を図1～図18を用いて説明する。なお、図中同一部分には同一符号を付して説明を省略する。

【0021】（音声記録媒体に係る第1の実施態様）まず、この発明に係る音声記録媒体に記録されるべき音声情報列の基本的な構造を図1を用いて概念的に説明する。

【0022】この発明に係る音声記録媒体で記録される情報は、映画における出演者の会話、日常の生活環境における会話等のように、長さの異なる複数のセンテンス（文）から構成され、また、各センテンス（各会話者の音声情報）の間に、音声再生されていない状況、雑音のみが再生されている状況、音楽（BGM）のみが再生

19

されている状況等のランダムに発生する無音声期間が存在し得る一連の音声情報列である。したがって、この音声記録媒体に係る第1の実施態様（基本構成）は、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長の区分（以下、セグメントという）に区分された音声情報列が記録された第1領域を少なくとも備えている。

【0023】一般にネイティブスピーカーの英会話では、1センテンスは概ね3秒程度で発声されるため、記録されるべき音声情報列を構成するセグメントを決定する発音の節目を各センテンスの間に設定することで、図1(a)、(b)あるいは(d)に示されたように、音声情報列を構成する可変長セグメントをそれぞれ構成するのが妥当である。なお、会話中のセンテンスの中には図1(c)に示されたように、極端に短いセンテンスも含まれるが、このセンテンスも1つのセグメントを構成する。一方、図1(e)に示されたように、極端に長いセンテンスの場合には、接続詞や関係詞等の前が発音の節目となるため、図1(e)に示されたようなセンテンスでは、連続する2つのセグメントで構成するのが妥当である。したがって、記録されるべき音声情報列のセグメントとは、発声上の区切り（息継ぎ位置）又は言語上（文法上）のなんらかの区切りにもとづいて分割された音声情報の記録単位であることを意味する。

【0024】また、当該音声記録媒体は、上述のような特殊な音声情報の再生中に発生する可能性のある割込み要求を適切に処理すべく、上述のように発音の節目で分割された各セグメントの前半部分と後半部分の境界を示す境界位置識別情報が記録された第2領域をさらに備えている。具体的にこの境界位置識別情報は、例えば当該音声記録媒体のディレクトリ領域や各セグメントのヘッダ部に記録される。

【0025】分割された各セグメントごとに設定される境界は、図1(a)、(b)あるいは(d)に示されたように、何等かの発音のまとまりがある最初の区切りS1、S2、S4に設定される（すなわち、S1、S2、S3、S4は各区分ごとに発音状態に応じて設けられ、これにより各区分は前半部分と後半部分に区切られる）。なお、図1(c)に示されたように、センテンスが極端に短い場合には、境界位置S3を分割されたセグメントの終了位置と一致するよう設定してもよい。また、図1(e)に示されたように、1つのセンテンスが複数のセグメントに分割された場合も、各セグメントについて発音のまとまりと判断された前半部分と後半部分とに区切る境界を示す境界位置S5、S6をそれぞれ設定する。

【0026】特に、この第1の実施態様に係る音声記録媒体の音声再生は、記録されたセグメントごとに順次に

(5)

特許2983194

11

行われる。しかしながら、係る音声再生中にリピート指示等の操作者からの割込み要求があった場合には、一旦この割込み要求の発生タイミングが判断され、この判断結果に基づいて該割込み要求の対象となるセグメントが特定される。具体的には、この割込み要求が、再生中であったセグメントの先頭位置から上述のように設定された境界位置までの前半部分で発生したか、あるいは該再生中であったセグメントの境界位置から次に再生されるべき次セグメントの先頭位置までの後半部分で発生したかが判断される。ここで、上記割込み要求が再生中であったセグメントの前半部分で発生したと判断されると、割込み要求が対象とするセグメントとして再生中であったセグメントの前のセグメント（既に音声再生が終了したセグメント）が特定される。一方、上記割込み要求が再生中であったセグメント後半部分で発生したと判断されると、割込み要求が対象とするセグメントとして再生中であったセグメント自身が特定される。

【0027】この再生方法は、このように特定されたセグメント（操作者の指示対象として特定されたセグメント）の先頭から、音声再生を再開することによりリピート機能を再現する。

【0028】（音声記録媒体に係る第2の実施態様）次に、この発明に係る音声記録媒体の第2の実施態様は、少なくとも2種類の音声情報列が記録されていることを特徴としている。すなわち、第1音声情報列は例えばネイティブスピーカが自然な速さで話す英語の音声情報からなり、この音声情報列は上述されたように発音の節目（センテンスの終りやセンテンス中の一息つける、発生上あるいは文法上の区切り）で複数の可変長セグメントに分割されている。第2音声情報列は上記第1音声情報列の内容と等価な意味内容であるが別の音声情報であり、例えば単語を区切って話すゆっくりとした速さの英語の音声情報である。なお、この第2音声情報列も、複数の可変長セグメントから構成されている。

【0029】ここで重要なことは、上記第1及び第2音声情報列はそれぞれ複数の可変長セグメントに区分されているが、互いにセグメントごとにその意味内容が対応していることである。例えば、第1音声情報列の1番目（図1(a)では621番目）のセグメントがネイティブスピーカの話す "It's not much of a problem." であるときは、第2音声情報列の1番目のセグメントは各単語を区切って話す "It is not much of a problem." となる。ただし、第2音声情報列と対応した内容でかつ別の音声情報からなるということは、言語上は同一の意味で発音の異なるものであることを示している。

【0030】さらに、この第2の実施態様に係る音声記録媒体は、所定領域に、記録位置識別情報が記録されていることを特徴としている。この記録位置識別情報は、少なくとも、第1及び第2音声情報列における各セグメントが、当該音声記録媒体のどの位置に記録されている

12

かを示す情報である。したがって、例えば第1音声情報列の1番目のセグメント "It's not much of a problem." に対応する第2音声情報列のセグメント "It is not much of a problem." が、媒体のどの位置に記録されているかということは、この記録位置識別情報により認識することができる。

【0031】その結果、第1及び第2音声情報列と記録位置識別情報は互いに無関係に記録されるのではなく、一定の関係をもって記録され、各音声情報列はセグメントを単位として有機的に組み合わされている。すなわち、第1及び第2音声情報列は互いに対をなしており、これらをセグメントごとに関連させているのが記録位置識別情報である。なお、この実施態様では、記録位置識別情報は当該音声記録媒体のディレクトリ領域に記録されており、少なくとも各セグメントの先頭位置に関する情報を含んでいる。また、この実施態様では境界位置識別情報は、各セグメントのヘッダ部分に記録される。

【0032】以上のような構造を備えた音声記録媒体（第2の実施態様）の再生方法でも、記録されたセグメントごとに順番に音声再生が行われるが、特に、この再生方法では、当該音声記録媒体に記録された第1音声情報列から第2音声情報列への再生切換え（あるいは第2音声情報列から第1音声情報列への再生切換え）が可能であることを特徴としている。なお、この再生切換え動作は、セグメントを単位として行われる。例えば、第1音声情報列の1番目のセグメントが再生されているときに第2音声情報列の再生指示が入力されると（割込み要求の発生）、記録位置識別情報に基づいて第2音声情報列の対応する1番目のセグメントを読み出し、その対応するセグメントの音声再生が実行される。また逆に、第2音声情報列から第1音声情報列への再生切換えも、上述した再生切換え動作と同様に各セグメント単位で行われる。

【0033】なお、上述の切換え動作において、切換え対象となるセグメントの特定は、割込みタイミングを上述された第1の実施態様に係る音声記録媒体の再生方法で説明されたように判断することにより行われる。また、この再生方法では、上述の再生切換え動作の他、リピート再生等の種々の変形が可能である。その代表的なものとして、いわゆる戻し指令がある。すなわち、再生中の停止命令により一時再生を中断した後に戻し指令が入力されたときは、指令された直だけ音声情報の読み出し位置を戻すことによりより操作者の希望に合った音声情報の再生が行われる。

【0034】（音声記録媒体の第3の実施態様）さらに、この発明に係る音声記録媒体の第3の実施態様について説明する。この第3の実施態様に係る音声記録媒体は、第1及び第2音声情報列の他、さらに第3音声情報列が当該音声記録媒体に記録されている点が、上記第2



13

の実施態様に係る音声記録媒体と異なる。

【0035】ここで重要なことは、上記第3音声情報列は第1及び第2音声情報列の1又は2以上の可変長セグメントをひとまとまりとしたセグメント群に区分されていることである。換言すれば、この第3音声情報列の1つのセグメント群は第1及び第2音声情報列の1又は2以上のセグメントを包含しており、したがって、第3音声情報列の1つのセグメント群は第1及び第2音声情報列の1又は2以上のセグメントと対になっている。特に、この構成は図1(e)に示されたように、1つのセ

【0036】また、この第3の実施態様の音声記録媒体において、所定の領域に記録された記録位置識別情報には、上記第3音声情報列の内容の記録位置をもセグメント群ごとに示す情報も含まれている。したがって、第1、第2及び第3音声情報列と記録位置識別情報は互いに一定の関係をもって媒体に記録され、各音声情報列はセグメントあるいはセグメント群を単位として有機的に組み合わせられている。なお、この第3の実施態様におい

【0037】以上のような構造を備えた音声記録媒体（第3の実施態様）の再生方法は、基本的に上述された第2の実施態様の場合と同じであるが、第1及び第2音声情報列間の再生切換えの他、該第1及び第2音声情報列と第3音声情報列との間においても再生切換え動作を行う点が、第2の実施態様の場合と異なる。

【0038】例えば、第1音声情報列の再生中にネイティブスピーカの“It's not much of a problem.”が聴き取れなかったときは、再生中の第1音声情報列から第2音声情報列に再生を切換えることにより、ゆっくりと単語を区切って話す“It is not much of a problem.”を聴くことができる。そして、この日本語の意味や文法を知りたいときは、さらに、再生中の音声情報列から第3音声情報列へ再生を切換えればよい。もちろん、この再生方法においても、上記第2の実施態様に係る音声記録媒体の再生方法で説明された展し指令や停止命令を組み合わせて使えるよう応用できることは言うまでもない。また、この再生方法においても、切換え再生及びリピート再生が可能であり、これら割込み要求が対象とするセグメントの特定は、上述された第1の実施態様の場合と同様に行われる。

【0039】（音声記録媒体の第4の実施態様）この発明に係る音声記録媒体の第4の実施態様は、基本的に上記第3の実施態様の場合と同様であるが、第1及び第2音声情報列の他、文字情報列が記録されている点が主に

(7)

特許2983194

14

異なる。この文字情報列は、第1又は第2音声情報列に対応する内容の文字情報に相当しており、例えばネイティブスピーカが話す英語（音声）に対応する文字情報に相当している。

【0040】この文字情報列も、第1及び第2音声情報列の各セグメントと対応するセグメントに区分されている。また、この第4の実施態様に係る音声記録媒体においても、記録位置識別情報には、この文字情報列の記録位置を各音声情報列のそれぞれのセグメントごとにそれらの先頭位置に関する情報が含まれ、当該音声記録媒体のディレクトリ領域に記録される。また、各セグメントのヘッダ部には境界位置識別情報が記録される。したがって、第1及び第2音声情報列と文字情報列はそれぞれセグメント単位で対応することになる。なお、この第4の実施態様に係る音声記録媒体において、上述の第3の実施態様における第3音声情報列を記録情報として加えるときは、第1及び第2音声情報列と文字情報列の1又は2以上のセグメントは第3音声情報列の1つのセグメント群にも対応することになる。さらに、この第4の実施態様においても、上記記録位置識別情報には、各セグメントの先頭位置が含まれ、かつ当該音声記録媒体のディレクトリ領域に記録される。また、境界位置識別情報は各セグメントのヘッダ部に記録される。

【0041】以上のような構造を備えた音声記録媒体（第4の実施態様）の再生方法も、基本的に上記第2の実施態様の場合と同様であるが、第1又は第2音声情報列の再生中に文字情報列がディスプレイ表示される点と異なる。

【0042】例えば、第1音声情報列のセグメント“It's not much of a problem.”が再生されているときは、所定の表示部に“It's not much of a problem.”もしくは“It is not much of a problem.”がディスプレイ表示される。なお、この表示については再生中の音声情報列と時間的に完全に同期している必要はなく、文字が少しずつ遅れて表示されたり、あるいは少しずつ先に表示されたりしてもよい。また、この再生方法においても、切換え再生及びリピート再生が可能であり、これら割込み要求が対象とするセグメントの特定は、上述された第1の実施態様に係る音声記録媒体の再生方法と同様に行われる。

【0043】次に、この発明に係る音声記録媒体の具体的な構造を、図2～図5を用いて、以下詳細に説明する。

【0044】図2は、この本発明に係る音声記録媒体の例として、上述の第3の実施態様を英会話練習用に適用したときの各音声情報列A、B、Cと、その記録内容を説明するための図である。この図において、音声情報列Aはネイティブスピーカの話す英語の情報列（第1音声情報列）であり、複数のセグメント621、622から構成されている。音声情報列Bはゆっくりと単語を区切

15

って話す英語の情報列（第2音声情報列）であり、これも先の音声情報列Aのセグメント621、622に対応する英単語あるいは句からなる複数のセグメントから構成されている。また、音声情報列Cは日本語の解説をする情報列（第3音声情報列）であり、この音声情報列Cに含まれるセグメント群は、各音声情報列A、Bの各セグメント621、622にそれぞれ対応している。

【0045】また、図3は、図2に示された態様における1セグメント当りの時間と容量の関係を説明するための表である。この表において、1秒間は6キロバイトの容量に対応している。例えば音声情報列Aのセグメント621では、“It's”の発声時間が0.2秒、その容量が1.2KB（キロバイト）、“not”の発声時間が0.1秒、その容量が0.6KB（キロバイト）、“much of a”の発声時間が0.4秒、その容量が2.4KB（キロバイト）、そして“problem”の発声時間が0.3秒、その容量が1.8KB（キロバイト）であり、セグメント621全体の発声時間は2.0秒、その容量は12KB（キロバイト）となる。なお、表中におけるカッコの数字は、境界位置を再生時間で示した値である。すなわち、この音声情報列Aの場合は、図1

(a)に示されたように、“not”の後に境界位置が設定されるため、セグメント621の再生開始から0.3秒後の位置に音声情報列Aを前半部分と後半部分に分ける境界が設定されていることを示している。また、他の音声情報列B、Cについても同様に示されているが、この図3の実施態様では、音声情報列Cには境界位置識別情報は設定されていない。

【0046】さらに、図4は、図2及び図3に示された態様におけるディレクトリ領域の記録内容を説明するための表である。この表において、ディレクトリ領域は、1セグメント当り $9 \times 3 = 27$ バイト（B）で構成される。音声情報列A、B、Cはそれぞれ図2の音声情報列A、B、Cに対応している。また、1バイトのCは属性を示し、C=0は音声情報列A、C=64は音声情報列Bであることを意味する。また、C=128、129は音声情報列Cであることを意味し、特にC=129のとき、すなわちビット表現（8ビット（bit））で“10000001”のときは前のセグメントと同じ解説対象であることを示す（音声情報列Cの解説対象となる同じセグメント群に属していることを示し、例えば図1

(e)のセグメント801、802の場合が相当する）。

【0047】位置情報のM、S、B（各1バイト）は産業界で標準になっているCD-ROM上の位置を表すパラメータである。すなわちMは分、Sは秒、Bはブロックをそれぞれ示す。また、1ブロックは2,048バイトであり、75ブロックで1秒分を構成している。したがって、最大の数はM=59、S=59、B=74となる。次の2バイトのSBはスタートバイトを示し、そ

(8)

特許2983194

16

の次の3バイトのLLLは各セグメント全体の長さを示している。なお、位置を示すパラメータに分、秒を使う理由はCD-ROMはもともと音楽用として開発されたためであり、始めからの時間として記録位置を表現するようになっている。そのためCD-ROMを当該音声記録媒体として採用した場合には、この分と秒は再生時の時間とは全く無関係であり、単に記録媒体上の記録位置を表わしている情報にすぎないことになる。

【0048】その結果、例えば音声情報列Aにおけるセグメント621の“It's not much of a problem.”は、0分11秒3ブロックの826バイト目から6,000バイトの長さでネイティブスピーカの話す英語の音声情報が記録され、音声情報列Bにおける対応するセグメントは0分11秒3ブロックの2,026バイト目から17,400バイトの長さでゆっくりと単語を区切って話す英語が記録され、音声情報列Cのセグメント群は0分11秒6ブロックの1,282バイト目から72,000バイトの長さで日本語解説が記録される。なお、621、622等のセグメントナンバーはメモリ上ではなく、そのアドレスに対応している。また、各セグメントの関係を示す記録位置識別情報は、このディレクトリ領域に含まれる。

【0049】さらに具体的には、第4図に示されたディレクトリ領域の記録内容から、当該音声記録媒体の0分11秒3ブロックにおける826バイト目から826+6,000-1=6,825バイト目までの領域には、セグメントが621で属性Cが0の音声情報列すなわちネイティブスピーカが話す“It's not much of a problem.”に相当する情報が記録される。また、当該音声記録媒体の0分11秒3ブロックにおける2,026バイト目から2,026+17,400-1=19,425バイト目までの領域には、セグメントが621で属性Cが64の音声情報列すなわち単語を区切ってゆっくり話す“It is not much of a problem.”に相当する情報が記録される。さらに、当該音声記録媒体の0分11秒6ブロックにおける1,282バイト目から1,282+72,000-1=73,281バイト目までの領域には、セグメントが621で属性Cが128の音声情報列すなわち日本語の解説に相当する情報が記録される。

【0050】このように、図4に示されたディレクトリ領域を設ければ、図3に示されたような再生時間及び容量で図2に示された各音声情報列が記録可能である。

【0051】次に、各セグメント621、622の前半部分と後半部分の境界を示す境界位置識別情報は、例えば図5(a)に示された可変長セグメントのヘッダ部に記録される。このヘッダ部は、図5(b)に示されたように、先頭から文字情報や画像情報の有無等を示すための1バイト領域（1B）、音声情報列A用に用意された領域であって情報列タイプ（音声情報列A、B等を区別するための情報）を示す1バイトデータ、そのデータ長



(9)

特許2983194

17

を示す3バイトデータ、及び境界位置を示す1バイトデータから構成された5バイト領域(5B)、音声情報列B用に用意された領域であって情報列タイプを示す1バイトデータ、そのデータ長を示す3バイトデータ、及び境界位置を示す1バイトデータから構成された5バイト領域(5B)、音声情報列C用に用意された領域であって情報列タイプを示す1バイトデータ及びそのデータ長を示す3バイトデータから構成された4バイト領域(4B)、文字情報列D用に用意された領域であって情報列タイプを示す1バイトデータ及びそのデータ長を示す3バイトデータから構成された4バイト領域(4B)、同様に文字情報列D用に用意された領域であってアドレスを示す3バイトデータ及びそのデータ長を示す3バイトデータから構成された6バイト領域(6B)、他の情報列(タイプE)用に用意された4バイト領域(4B)、及び予備の3バイト領域(3B)からなる、32バイトの領域である。

【0052】なお、この明細書中で説明される実施態様では、他の情報列(タイプE)は使用していない。また、この実施態様では、境界位置識別情報は音声情報列A、Bにのみ設定されている。境界位置を示す1バイト領域には、圧縮後の情報においてセグメント先頭からのバイト数を256で割った商がセットされる。例えばAトラック(音声情報列Aに相当する仮想トラック)の音声サンプリングレートを128キロビット/秒(16キロバイト/秒)とすると、この1バイト領域で先頭から4秒までの範囲の境界位置が表せ、その分解能は0.016秒となる。また、音声サンプリングレートが32キロビット/秒(4キロバイト/秒)の場合には、この1バイト領域で先頭から16秒までの範囲の境界位置が表せ、かつその分解能は0.064秒となり、実用上十分な性能が得られる。

【0053】次に、図6～図10を用いて、この発明に係る音声記録媒体の基本的な再生方法を説明する。

【0054】まず、図6は、この発明に係る音声記録媒体の再生方法実現するための再生装置の全体構成を示す斜視図である。この図からも分かるように、当該音声記録媒体は、例えばポータブルなCDプレイヤー(再生装置本体2)により再生可能なCD-ROMであり、この再生装置本体2はコード接続されたハンドセット8によりリモート制御される。このハンドセット8には少なくとも再生中のセグメント番号を表示する液晶ディスプレイ(LCD)等の表示部21や、各種制御用ボタン群240が設けられている。また、操作者は再生装置本体2で再生された音声情報をイヤホン13を介して聴くことができる。

【0055】また、図7は、図6に示された再生装置の構成を示すブロック図である。この図に示されたように、当該音声記録媒体1であるCD-ROMは再生機構20にセットされる。再生機構20はディスクインター

18

フェイス(I/F)3及びバス4を介してCPU5に接続されている。また、バス4にはプログラムを格納するための例えば32キロバイト(KB)のROM6と、ディレクトリや音声情報列を一時的に格納するための例えば256キロバイトのRAM7とが接続されている。さらに、バス4には手動操作のためのハンドセット8との間で情報の授受を行なうハンドセットインターフェイス(I/F)9と、音声出力用のアンプ(AMP)10を介して外部端子11及びハンドセット8に接続されたD/Aコンバータ12に接続されている。なお、ハンドセット8には上述されたようにイヤホン13が接続されている。

【0056】図8(a)、(b)は、それぞれROM6及びRAM7のメモリ割り当て状況を説明するための図である。図8(a)に示されたように、32キロバイトのROM6にはプログラムが格納される。一方、図8(b)に示されたように、RAM7には、 $(50+50)=100$ キロバイトのバッファ(50ブロック分に相当)と、 $(75+75)=150$ キロバイトのディレクトリと、6キロバイト分のシステムエリアが割り当てられる。したがって、RAM7には常時50ブロック分の音声情報列が保持され、かつ150キロバイト+27 $\times$ 5、555セグメント分のディレクトリ(音声情報列Aの部分のみで約30分間に相当)が保持される。

【0057】なお、上述の具体例では当該音声記録媒体としてCD-ROMを用いているが、その代表的なものの容量は552メガバイト(MB)である。CD-ROMではアドレスを表わすのに分、秒、ブロックの単位を用いている。また、1ブロックは2,048バイト、75ブロックは1秒、60秒は1分であるため、該CD-ROMのアドレスの最大の値は59分59秒74ブロックである。逆に、このCD-ROMの容量は2,048 $\times$ 75 $\times$ 60 $\times$ 60=552,96メガバイトである。このうち、最初から2秒分はCD-ROMのフォーマットとしてユーザは使えないので、正値には最大容量とし552,6528MBとなる。さらに、最初から20秒に相当するところまでディレクトリ領域が割り当てられると、3メガバイトのディレクトリ容量をCD-ROMに確保することができる。

【0058】次に、容量に関する計算例を示す。

【0059】音声のサンプリングレートはADPCM方式による16キロサンプリング/秒とし、1サンプリング当たり3ビットとする。このように仮定すると音声サンプリングは48キロビット/秒(Kbit/s)すなわち6キロバイト/秒(KB/s)となり、転送レートもこれに合わせる必要がある。なお、16キロサンプリング/秒であれば8キロヘルツ(KHz)までのf特性があり、子音まで十分に記録できる。このように仮定すると、1時間の録音に6キロバイト $\times$ 3,600秒=21,6メガバイトの容量が必要となる。一般に、CD-ROM1枚には

(10)

19

エラーコレクションを入れて552メガバイト記録できる。ディレクトリ領域を除くと549メガバイトを各音声情報記録されるデータ領域として使用できる。したがって、 $549 \div 21.6 = 25$ 時間24分の音声情報を記録することが可能となる。そこで、英会話用として用いる場合には、ネイティブスピーカによる自然な速度で発音されている物語を1時間とすると、それを単語1つ1つを区切ってゆっくり発音している部分はその4倍で約4時間とみることができる。そして、解説の部分は合計で15時間あったとしても全部で20時間である。

【0060】次に、1時間の会話部分をいくつかのセグメント(区画)に分解できるかを考えると、1秒間に平均4個の区切りとして1時間で14,400個である。1個のセグメントのディレクトリには36バイト必要なので全体で約518キロバイトであり、これは上記3メガバイトのディレクトリの収納場所に充分収納でき、かつ1時間の物語の全てのディレクトリを格納することができる。

【0061】次に、図9は、図6及び図7に示されたハンドセット8とイヤホン13の構成を具体的に示す平面図である。ハンドセット8の前面にはセグメント番号を表示する表示部21(LCD)、現時点の指示内容をLEDを利用して示す表示部22、再生モード(リピート再生が実施されている場合等の再生状態を示す)を切替えるためモードスイッチ81、各音声情報列A、B、Cの再生指令をするための指令スイッチ23A、23B、23C、機能指示するためのADVボタン24A、再生の戻し指令をするためのREVボタン24B、リピート再生を指示するためのREPボタン25A、再生の停止を指令するためのSTOPボタン25B、及びイヤホン13から出力される音量を調節するためのボリュームスイッチ28が設けられている。また、イヤホン13はコード26を介してハンドセット8に接続され、ハンドセット8はコード27を介して再生装置本体2に接続されている。

【0062】次に、この発明に係る音声記録媒体の再生方法における、基本的な再生シーケンスの一例を説明する。図10は、この基本的な再生シーケンスを説明するための図である。音声情報列Aの部分すなわちネイティブスピーカの英語を続けて聴くときは、この図10(a)に示されたように順次記録されたセグメントの音声情報が再生される。この場合は、図9に示されたボタン23Aを押してそのまま聴いているだけでよい。

【0063】一方、図10(b)に示された再生シーケンスでは、まず、セグメント621から順に音声情報列Aの再生が行われる。ここで、例えばセグメント623がよく聴き取れなかった時には、操作者はすぐにSTOPボタン25Bを押すことにより一旦音声再生を中断する。このとき、セグメント表示部21のセグメント番号は624になっている。そこで、REVボタン24Bを

20

特許2983194

1回だけ押す。なお、REVボタン24Bは押しつづけていると連続してもとり、1回押すごとに表示部21に表示されたセグメント番号は1つずつ戻る。セグメントが624の状態ではREVボタン24Bが1回押されると、表示部21に表示されたセグメント番号は623になる。

【0064】続いて、スイッチ23Bが押されると(切換え再生指示)、音声情報列Aのセグメント623に対応した音声情報列Bのセグメント623が再生され、1つ1つの単語を区切ってゆっくり発音している音声が届える。そのままにしておくと続けて音声情報列Bのままセグメント624の再生動作が進行する。さらに、この音声情報列Bのセグメント624を聴いたところでSTOPボタン25Bが押された後、再度REVボタン24Bが押されると表示部21のセグメント番号は623になる。この状態でスイッチ23Cが押されると、音声情報列Cの日本語の解説が再生される。解説の部分はいくつかの単語がまとまった句について解説していることもあるので、いくつかのセグメント番号で一つの解説ということもある属性Cを示すバイトの最下位ビットが1のもの(属性Cが129のもの)はそれを示し、同一の音声情報列の前のセグメント番号と同じ解説対象であることを意味している。ここに当たったら読みとばすこともできる。

【0065】次に、図11～図13を用いて、この発明に係る音声記録媒体の他の実施態様を説明する。

【0066】図11は、この発明に係る音声記録媒体におけるディレクトリ領域の他の構成例を示す図である。この図に示されたディレクトリは、1セグメント当たり15バイトで構成されている。すなわち、1バイトの属性Cと、2バイトのセグメントナンバーSSと、1バイトのサフィックスナンバーNと、1バイトの分Mと、1バイトの秒Sと、1バイトのブロックBと、そして3バイトのセグメント長LLLの計12バイトである。

【0067】また、属性Cの8ビットにおいて、第1ビット(最上位ビット)が1のときはセグメントのスタートを意味し、0のときはその他の状態であることを意味する。また、この属性Cにおける第2、3ビットで表現される値が0のときは音声情報列Aであることを意味し、1のときは音声情報列Bであることを意味し、2のときは音声情報列Cであることを意味する。さらに、第4ビットが0のときはサフィックスが無いことを意味し、1のときはサフィックスが有ることを意味する。第5ビットが1のときは前のセグメントと同じセグメント群に関連づけられていることを意味し、0のときは同じセグメント群には関連づけられていないことを意味する。

【0068】図12は、図11に示されたディレクトリに対応した各音声情報列A、B、Cのセグメントの分割状況を説明するための図である。そして、これが図2に

(11)

特許2983194

21

示された例と異なる点は、サフィックスを付したセグメントで区分されている点である。

【0069】また、図13は、以上説明されたように、8ビットの属性Cの第2、3ビットの値と分割された各図に相当するセグメントとの関係を示す図である。

【0070】さらに、図14及び図15を用いて、この発明に係る音声記録媒体の他の実施態様の説明しておく。なお、この他の実施態様が先に説明した図11～図13の実施態様と異なる点は、この実施態様が音声情報だけでなく文字情報の記録、再生をも行なうようにしている点である。

【0071】第14図はそのディレクトリの構造を説明するための図である。この図において、先に説明した図11のディレクトリと異なる点は、属性Cの第2、3のビットに文字情報Dに関する情報が付加される点である。このビットが3であるときは、文字情報Dが当該音声記録媒体に所定のコードで記録されていることを意味する。

【0072】図16は、この実施態様に係る音声記録媒体の再生装置に用いられるハンドセットの構成を示す平面図である。なお、この図15のハンドセットと図9のハンドセットとの差異は、文字情報Dについても表示可能な態様である点である。具体的に、ハンドセット8には、文字表示を指示するための操作ボタン23Dとともに、LCDで構成される文字表示部210が設けられている点である。また、この実施態様では、文字の表示をオン、オフさせるための文字表示ボタンもさらに設けられている。このようなハンドセット8では、文字表示部210で文字情報Dを表示するか否かの設定が可能である。

【0073】次に、図11～図13に示された他の実施態様の作用の特色を説明する。

【0074】この実施態様では、音声情報列Aの再生中は情報列Aの発音のスピード、すなわちLLLの長さに応じてその文字情報列Dの現われるスピードも制御される。つまり、このセグメントの発音の始まりから終わりまでの間に、文字が表示部210に現われ始めて完全に出力し終るようになる（発音に完全に同期させて文字を表示させる）。

【0075】次に、音声情報列Bの再生中は音声情報列Aの部分の再生に比べて数倍の時間をかけて発音されているので、この時の音声情報列Bの発音の長さに同期させて文字情報列Dを表示させる。実際には、文字の表示の方を早めに出す（音声出力の方を少し遅らせて出す）方が使う人にとって便利なのも考えられるため、この点は任意に設定すればよい。

【0076】次に、図14及び図15に示された実施態様として、情報記録媒体をランダムアクセスのできない磁気テープとした場合を説明する。

【0077】この実施態様においても、基本的な情報の

22

記録、再生の態様は上述の実施態様と同様であるが、磁気テープではデータの飛び越し再生が容易でないため、実用的な再生装置とするためには、情報を一時的に記憶しておくためのバッファを設けるのが望ましい。具体的には、前述の仮想トラックの音声情報列A、B、C（3トラック）に相当する各小単位を、 $A_1, B_1, C_1, A_2, B_2, C_2, A_3, B_3, C_3, \dots$ というように、対応する音声情報列A、B、Cの各セグメントを1つのかたまりとして順番にならべる。そして、音声情報列Aを再生する時は $A_1, A_2, A_3$ というように音声情報列Aのセグメントのみ取り出して再生し、音声情報列Bを再生する時は $B_1, B_2, B_3, \dots$ というように音声情報列Bのセグメントのみ取り出して再生する。音声情報列Cについても同様である。この時、セグメント $A_1$ の次にセグメント $A_2$ を切れ目なく再生するためにはこれらセグメントの間に存在するセグメント $B_1, C_1$ を飛び越さなければならない。これは、通常のテープ走行では容易でないため、磁気テープからは早めにバッファの中に情報を取ってきておく。そして、飛び越して再生するのはバッファの中で行うようにする。

【0078】この磁気テープを、いわゆるDAT（デジタル・オーディオ・テープ）としたときには、次のようになる。まず、音声情報列Aのセグメントとして、Aトラックに記録されるネイティブ・スピーカの発音をある単位で区切る。例えば、ここでは1秒毎の一定時間に区切ることにする。そして、トラックB（音声情報列B）としては、ゆっくり（例えば平均1/3の速度で）話している内容のものとする。Aトラックの1秒に相当するところをBトラック用に4秒分とれるよう領域を確保する。そして、解説用のCトラックとしては9.4秒分とれるよう領域を確保する。なお、上記の秒数は一例であり、これに限定されるものではない。つまり、14.4秒を1つの単位とする。すなわち、14.4秒はA、B、Cトラック（各音声情報列A、B、Cに相当）の各対応する3個のセグメントをひとまとめたものである。ここで注意を必要とするは、Aトラック用の1秒分はオリジナルのものを1秒ずつ切っているだけで、もとの1秒と完全に一致していることである。Bトラックの4秒は4秒以内であればよく、3秒で終わって1秒あまってもよい。なお、この余ったところ（1秒分）を再生する時に飛ばす方法については後述する。

【0079】ロータリ・ヘッド型DATの場合は次のようになる。すなわち、このロータリ・ヘッド型DATでは、ロータリ・ヘッドの1回転で2トラックを記録又は再生する方式が一般的である。そして、1トラックに2,880バイトの音声情報を記録できるようになっている。したがって、30トラックを1区切りとすると、86,400バイトとなる。この中に、上述の実施態様でも説明された音声のサンプリング・レート=48(Kbit/s)で音声記録すると、

(12)

特許2983194

23

86,400(バイト)÷6,000(バイト)=14.4(s)

だけ記録できる。

【0080】次に、バッファについて考察する。

【0081】1MBのRAMを使用すると

10,480,576÷86,400=12,135

となるので、12個のバッファを設けることができる。

12個のバッファを機械的な意味でリング状にならべ、セグメント情報を常に2個ずつ先取りしてテープからバッファに転送する。再生はこのバッファに転送された情報を利用して行われる。任意のところで止めた時には、10個のバッファが止めた時点より前の音声(既に再生音としてスピーカから出してしまったもの)が残っている。すなわち、Aトラック上の音声にして10秒の音声

がテープを巻き戻すことなく、バッファから繰り返し再生することができる。

【0082】このようにすることにより、10個のバッファに残っている10秒間の音声情報については、前述したCD-ROMを用いた実施例と同じく、任意のセグメントにおいてA、B、Cのどの仮想トラックにも容易

に移ることができる。

【0083】ただし、CD-ROMのようにランダム・アクセスができないので、A、B、Cの3つのトラックの対応するセグメントの時間数が、常に14.4秒以内にならなければならない。この時、仮想トラックAの部分は、常に一定の時間幅(例えば1秒毎)になるように、機械的に区切られていてもよい。例えば、Bトラックの部分は2秒の時もあるし5秒の時もあるので、可変になるようにする。14.4秒からA、Bトラックの時間を引いた時間がCトラックに割り当てられる時間である。なお、B、Cトラックのいずれも、区切りの中のAトラックに完全に対応した部分を記録する必要はなく、近くの複数の区切りを通して、あるまとまりができていればよい。

【0084】DATにおいてはサブコード領域があるので、各区切り(14.4秒/30トラック毎の)と、各区切りの中のA、B、Cトラックの境目を示すバイト数の情報は、このサブコード領域に記録することができる。

【0085】次に、DATに記録可能な音声情報量を具体的に計算する。

【0086】一般に、1時間用のDATには240,000トラックが記録できる。30トラックを1区切りにしているが、次の6トラックを休みのトラックに使い、テープを止めている時にロータリーヘッドが多数回通っても、音声情報の記録されている部分にキズをつけないようにする。このようにするためには、36トラックが一単位となる。つまり、全体では6,666単位が記録でき、6,666×14.4秒=126時間39分50秒となる。このうち、Aトラック分として1時間に割り

24

当てるので、B、Cトラックに25時間39分50秒を使えることになる。ゆっくり発音したBトラックが4時間分使ったとしても、説明用のCトラックに21時間以上割り当てることができ、十分な長さが確保できる。

【0087】なお、上記のA、B、Cトラックのセグメントのとり方は、前述のCD-ROMを応用したものと同じくAトラックは1秒等一定時間幅で区切るだけでなく、秒程度を目安とし、何らかの発音上の切り目があるところで切ってもよい。すなわち、記録する時の編集方針により変化してもよい。したがって、CD-ROMのようにランダム・アクセスのできるものでも、一定時間幅で区切ってセグメントを区分することも可能である。

【0088】次に、この発明に係る音声記録媒体の再生方法、特にリピート再生要求や切替え再生に関する割込み処理を図16～図18を用いて説明する。

【0089】図16は、この発明に係る音声記録媒体の再生方法における、割込み対象セグメントの決定動作を説明するためのタイムチャートである。

【0090】通常、この発明に係る音声記録媒体の再生方法では、あらかじめ指示された音声情報列のセグメント(…、N-1、N、N+1、…)を順に再生していく。この再生途中で操作者(例えば、英語学習者等)からリピート再生や切替え再生等に関する割込み要求(図中、R1、R2、R3、…で示す)が発生したとき、まず、割込み要求の対象となったセグメントが特定される。すなわち、図16に示されたように、セグメントNが再生されているときに、係る割込み要求が先に説明された各セグメントのヘッダ部に記録された境界位置識別情報で示される境界位置 $S_n$ よりも前か後かが判断される。なお、図中において各セグメント内の境界位置は、それぞれ…、 $S_{n-1}$ 、 $S_n$ 、 $S_{n+1}$ 、…で示されている。さらに、この割込み要求の発生タイミングに関する判断結果に基づいて、当該割込み要求の対象セグメントとして再生中であったセグメントNかあるいはその前のセグメントN-1が特定される。発生した割込み要求の対象セグメントが特定されることにより、該割込み要求の指示内容にしたがって、この特定されたセグメントの先頭か、あるいは該特定されたセグメントに対応する他の音声情報列のセグメントの先頭から、音声情報の再生動作が再開される。

【0091】具体的に、例えばある音声情報列のセグメントNの再生中に所望の割込み要求R1が発生した場合、この割込み要求R1の発生タイミングはセグメントNの前半部分で発生したことになる。この場合は操作者の指示遅れと判断して再生中であったセグメントNの1つ前のセグメントN-1を割込み要求の対象セグメントとして特定する。一方、再生中であったセグメントNの後半部分で割込み要求R2が発生した場合は、この再生中であったセグメントN自身が割込み要求の対象セグメントとして特定される。なお、他のセグメント例えば図

(13)

特許2983194

25

16に示されたセグメントN+1の再生中に発生した割込み要求R3の場合も同様に、境界位置 $S_{n+1}$ を基準に割込み要求の発生タイミングを判断し、割込み要求の対象セグメントが特定される。

【0092】図17は、この発明に係る音声記録媒体の再生方法における、リピート再生割込み処理（再生方法に係る第1実施態様）を説明するためのフローチャートである。

【0093】この再生方法で用意される音声記録媒体には、少なくとも、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長のセグメントに区分された音声情報列と、該音声情報列における各区分ごとに、発音状態によって区切られた該各区分の前半部分と後半部分との境界を示す境界位置識別情報が、それぞれ所望の領域に記録されていけばよい。

【0094】まず、以上のような音声記録媒体の再生中において、リピート再生に関する割込み要求が発生すると、一旦再生中であったセグメントの情報（セグメントナンバー）が所定の保持手段によって保持されるとともに（ステップST1）、当該割込み要求の発生位置に関する情報も一旦保持される（ステップST2）。そして、再生中であったセグメントのヘッダ部に記録されていた境界位置と割込み発生位置とが時間軸に沿って比較される（ステップST3）。このとき、割込み要求が境界位置よりも後で発生した場合には、再生中であったセグメント自身が当該割込み要求の対象セグメントとして特定され、当該特定されたセグメントの先頭位置情報がディレクトリから読み出されて音声再生動作が再開される（ステップST4）。一方、割込み要求が境界位置よりも前で発生した場合には、再生中であったセグメントより1つ前のセグメントが当該割込み要求の対象セグメントとして特定され、当該特定されたセグメントの先頭位置情報がディレクトリから読み出されて音声再生動作が再開される（ステップST5）。

【0095】なお、この実施態様では、指示されたリピート再生動作は、次の新たな指示があるまで、特定されたセグメントの音声情報が繰り返して再生される。

【0096】次に、この発明に係る音声記録媒体の再生方法における、切替え再生割込み処理（再生方法に係る第2の実施態様）を、図18のフローチャートを用いて説明する。

【0097】この実施態様で用意される音声記録媒体は、所定の音声再生手段で再生出力されるべき複数の単語列から構成された1又は2以上の文に対応する音声情報列であって、発音の節目でそれぞれ分割された音声情報ごとに可変長セグメントに区分された音声情報列Aと、この音声情報列Aと等価でかつ遅延した発音で構成された音声情報列であって、音声情報列Aの各セグメン

26

トの音声情報と等価な音声情報ごとに可変長セグメントに分割された音声情報列Bと、音声情報列A、Bを切替えて再生すべく、切替え可能な各セグメントを当該記録媒体における該各セグメントの記録位置で示す記録位置識別情報と、そして、音声情報列A、Bにおける各セグメントごとに、発音状態に応じて区切られた（発音のまとまりと判断された）前半部分と後半部分との境界を示す境界位置識別情報とが、少なくとも記録されていけばよい。

【0098】まず、以上のような音声記録媒体の再生中において、切替え再生に関する割込み要求が発生すると、一旦再生中であったセグメントの情報（セグメントナンバー）が所定の保持手段によって保持されるとともに（ステップST6）、新たに指示された音声情報列の情報が保持される（ステップST7）。また、当該割込み要求の発生位置に関する情報も一旦保持される（ステップST8）。そして、再生中であったセグメントのヘッダ部に記録されていた境界位置と割込み発生位置とが時間軸に沿って比較される（ステップST9）。このとき、割込み要求が境界位置よりも後で発生した場合には、再生中であったセグメント自身が当該割込み要求の対象セグメントとして特定され、割込み要求で指示された音声情報列のセグメントであって、該特定されたセグメントの先頭位置情報がディレクトリから読み出されて音声再生動作が再開される（ステップST10）。一方、割込み要求が境界位置よりも前で発生した場合には、再生中であったセグメントより1つ前のセグメントが当該割込み要求の対象セグメントとして特定され、割込み要求で指示された音声情報列のセグメントであって、当該特定されたセグメントに対応するセグメントの先頭位置情報がディレクトリから読み出されて音声再生動作が再開される（ステップST11）。

【0099】なお、この発明は上述の各実施態様に限定されるものではなく、種々変形が可能である。

【0100】例えば、ビデオ・ディスクやビデオ・テープとの連動でもこの発明が適用できる。すなわち、普通のスピード（ネイティブ・スピーカの発音スピード）のAトラックに例えば映画のサウンド・トラック部を入れておき、Bトラックにゆっくりした発音を入れておき、Cトラックに日本語解説を入れておく。そして、Aトラックの再生を行ないながら、理解できなかったところで止め、少し戻し（この時、画面は固定のままでも実用に耐える）でBトラックを聴き、それでも分らなければCトラックを聴く。なお、Aトラックを再生する時だけ画面と同期させればよい。

【0101】また、パソコンなどにおいてこの発明を用いることもできる。すなわち、CAI（Computer Aided Instruction）との連動やCDI（Compact Disk Interactive）等高度のプログラマブルなものとも結合すれ

(14)

特許2983194

27

ば、より高度の応用が可能である。また、パソコンの記憶装置を上記のCDやDATの代りに用いてこの発明を実施することも可能である。

【0102】さらに、英会話などに限らず、詩吟、法律用のものにも適用することができ、情報列は3種又は4種のものに限らず、それ以上であってもよい。

【0103】

【発明の効果】以上のようにこの発明によれば、各種音声情報とともに、該音声情報列を構成する可変長セグメントに関する情報として、記録媒体における該各セグメントの発音状態に応じて区切られた前半部分と後半部分との境界を示す境界位置識別情報が少なくとも記録された音声記録媒体を用意し、所定の割込み要求のタイミングが再生中であったセグメントの境界位置の前か後かを判断し、該判断結果に基づいて割込み要求の対象セグメントを特定するしているので、操作者の指示遅れも考慮されたフレキシブルな再生動作が実現できるという効果がある。

【図面の簡単な説明】

【図1】この発明に係る音声記録媒体に記録されるべき音声情報列を含む各種情報を概念的に説明するための図である。

【図2】この発明に係る、英会話練習用に適用された音声記録媒体（第3の態様）の各音声情報列と、その記録内容を説明するための図である。

【図3】図2に示された各音声情報列について、各セグメント当りの時間と容量との関係を説明するための表である。

【図4】図2及び図3に示された音声記録媒体におけるディレクトリ領域の記録内容（記録位置識別情報を含む）を説明するための表である。

【図5】この発明に係る音声記録媒体に記録されるべき可変長セグメントの構成（境界位置識別情報を含む）を示す図である。

【図6】この発明に係る音声記録媒体の再生方法を表現する再生装置の全体構成を示す斜視図である。

【図7】図6に示された再生装置の構成を示すブロック図である。

28

\*【図8】図7に示されたROM及びRAMのメモリ割り当て状況を説明するための図である。

【図9】図6に示されたハンドセット及びイヤホンの全体構成を示す平面図である。

【図10】この発明に係る音声記録媒体の再生方法における基本的な再生シーケンスを説明するための図である。

【図11】この発明に係る音声記録媒体におけるディレクトリ領域の他の構成例を示す図である。

10 【図12】図11に示されたディレクトリ領域の記録内容に対応した各音声情報列のセグメントの分割状況を説明するための図である。

【図13】図12に示された分割状況において、ディレクトリ領域の記録内容のうち8ビットの属性Cにおける第2及び第3ビットの値と分割された各セグメントとの関係を説明するための図である。

【図14】この発明に係る音声記録媒体におけるディレクトリ領域の他の構成例を示す図である。

20 【図15】図12に示されたディレクトリを有する音声記録媒体の再生を行う再生装置の、ハンドセットの概略構成を示す平面図である。

【図16】この発明に係る音声記録媒体の再生方法における、割り込み対象セグメントの決定動作を説明するための図（タイムチャート）である。

【図17】この発明に係る音声記録媒体の再生方法における、リピート再生に関する割り込み処理を説明するためのフローチャートである。

【図18】この発明に係る音声記録媒体の再生方法における、切換え再生に関する割り込み処理を説明するためのフローチャートである。

30 【図19】従来の音声記録媒体の再生方法における、リピート割り込み動作を説明するための図（タイムチャート）である。

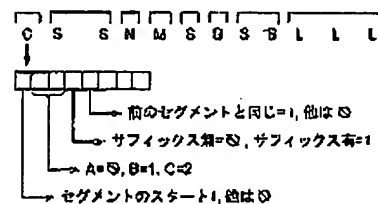
【符号の説明】

1…記録媒体、2…再生装置、8…ハンドセット、13…イヤホン、21、210…表示部、23A、23B、23C、23D…再生情報列指示スイッチ、25A…リピート指示スイッチ。

【図3】

セグメント		021	022
時間 (秒)	A	2.0 (0.3)	1.3 (0.4)
	B	3.0 (1.1)	2.6 (1.4)
	C	12	25
容量 (KB)	A	12	7.8
	B	23.4	15.6
	C	72	160

【図11】

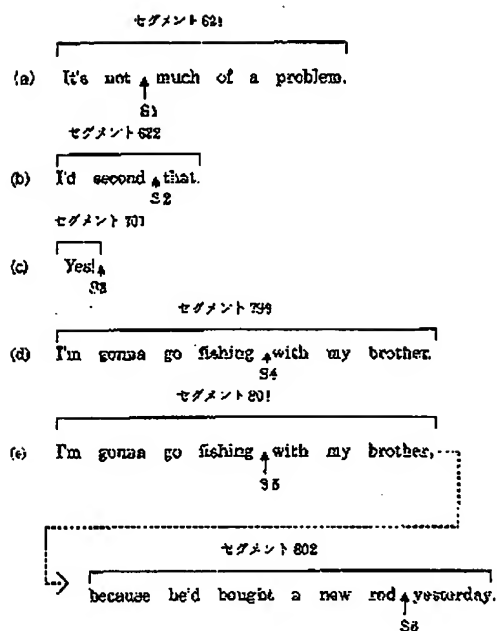




(15)

特許2983194

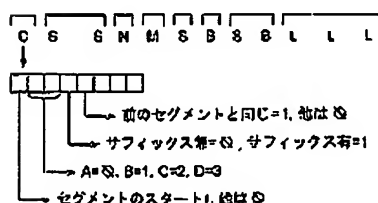
【図1】



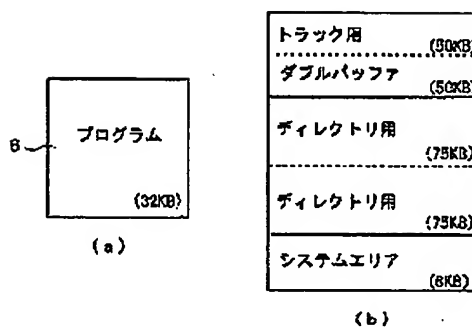
【図2】

記録内容	
A	<div> <div>"It's not much of a problem."</div> <div>"I'd second that."</div> </div> <div> <div>621</div> <div>622</div> </div>
O	<div> <div>"It is not much of a problem. I would second that."</div> </div> <div> <div>621</div> <div>622</div> </div>
C	<div> <div>(621)</div> <div>「I」 とは前の〇〇を指しており、「not much a problem」は動詞句で「強迫がない」とか「心配ない」というような意味です。</div> </div> <div> <div>(622)</div> <div>次に、「I'd」とは「I would」をつめた言い方であり、同じように「I could」も「I'd」とつめて言えます。</div> <div>「second」とは「支持する」、「賛成する」という意味があり、「I'd second that」で「同意です」という意味になります。</div> </div>

【図14】



【図8】



【図13】

2, 3ビット目	セグメント
0	621
1	621-1
1	621-2
2	621-1
2	621-2
2	621-3
2	621-4
0	622
1	622
2	622-1
2	622-2

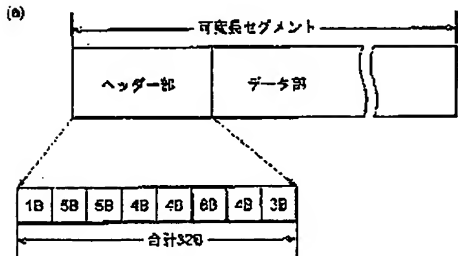
(15)

特許2983194

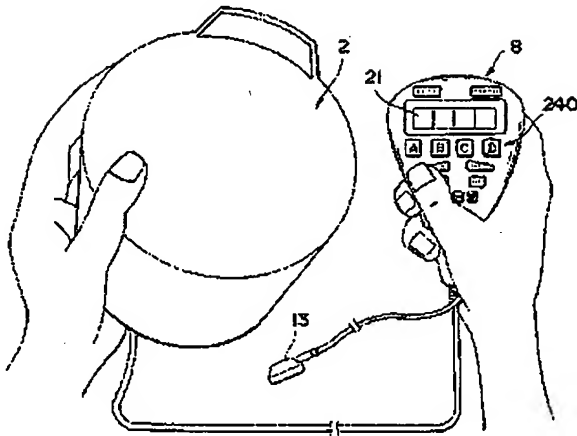
【図4】

情報列	セグメント SS	カテゴリ C	位置情報 M S B			スタートバイト SB	セグメント量 (バイト) LLL
A							
A	621	0	0	11	3	028	8,000
B		64	0	11	8	2,026	17,400
C		128	0	11	6	1,282	72,000
A	622	0	0	11	49	2,018	4,800
B		64	0	11	50	570	12,800
C		128	0	11	52	1,274	150,000

【図5】



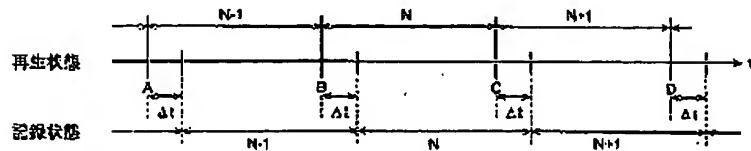
【図6】



(b)

ヘッダー (合計32B)	情報列	内容
1B		文字情報や図象情報の有無等 (1B)
5B	A	情報列タイプ (1B), データ長 (3B), 境界位置 (1B)
5B	B	情報列タイプ (1B), データ長 (3B), 境界位置 (1B)
4B	C	情報列タイプ (1B), データ長 (3B)
4B	(D)	情報列タイプ (1B), データ長 (3B)
6B	D	アドレス (3B), データ長 (3B)
4B	E	
3B	予備	

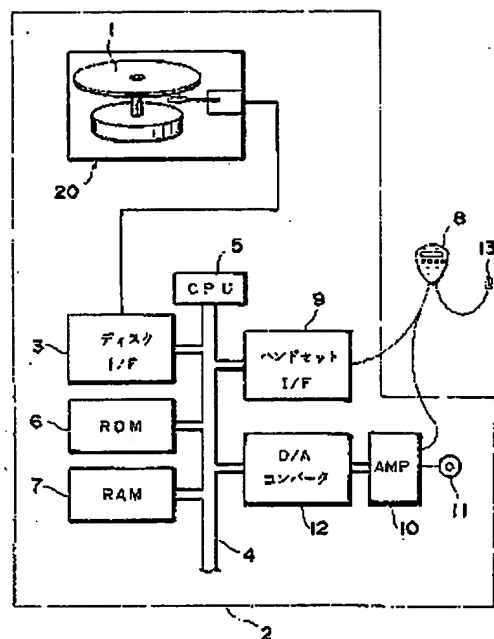
【図19】



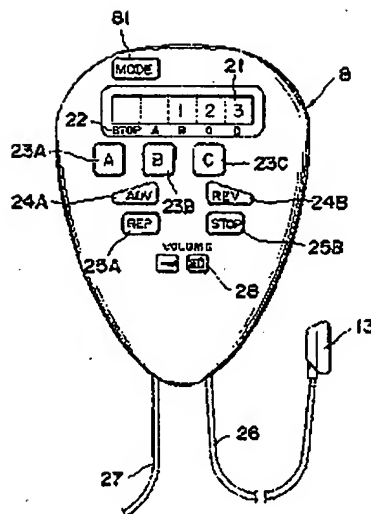
(17)

特許2983194

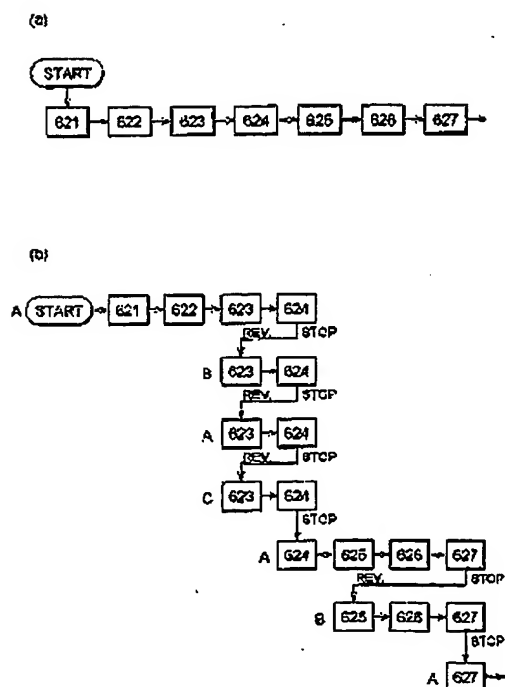
【図7】



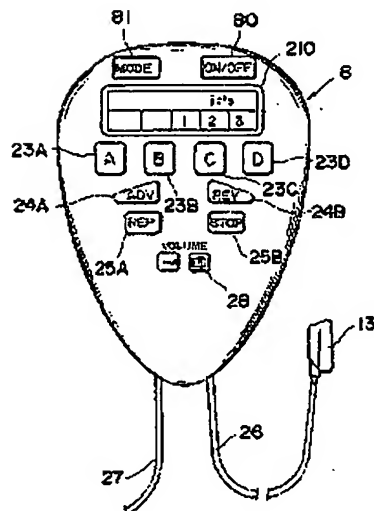
【図9】



【図10】



【図15】



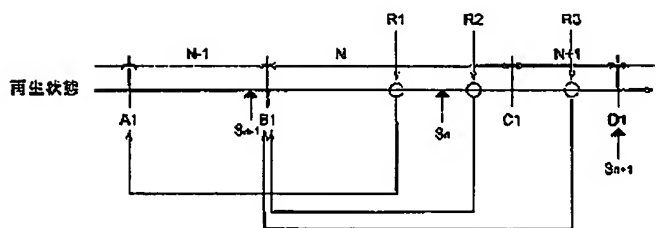
(18)

特許2983194

【図12】

情報列	セグメント分割
A	<div> <div>It's not much of a problem.</div> <div>I'd second that.</div> </div> <div> <div>621</div> <div>622</div> </div>
B	<div> <div>It's not much of a problem.</div> <div>I'd second that.</div> </div> <div> <div>621-1</div> <div>621-2</div> <div>622</div> </div>
C	<div> <div>621-1</div> <div>621-2</div> <div>621-3</div> <div>621-4</div> <div>622-1</div> <div>622-2</div> </div>

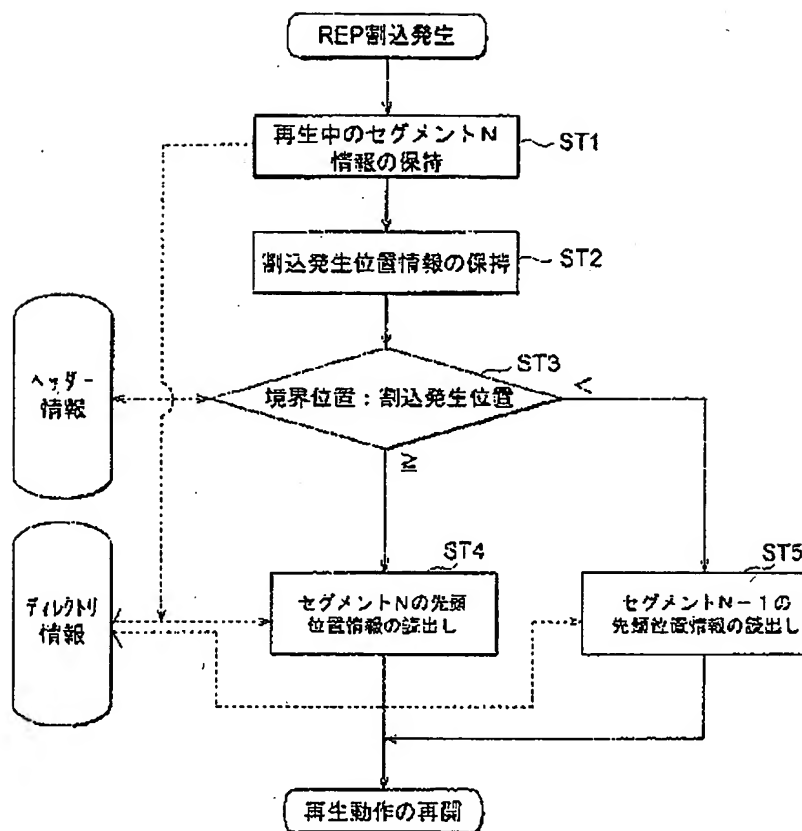
【図16】



(19)

特許2983194

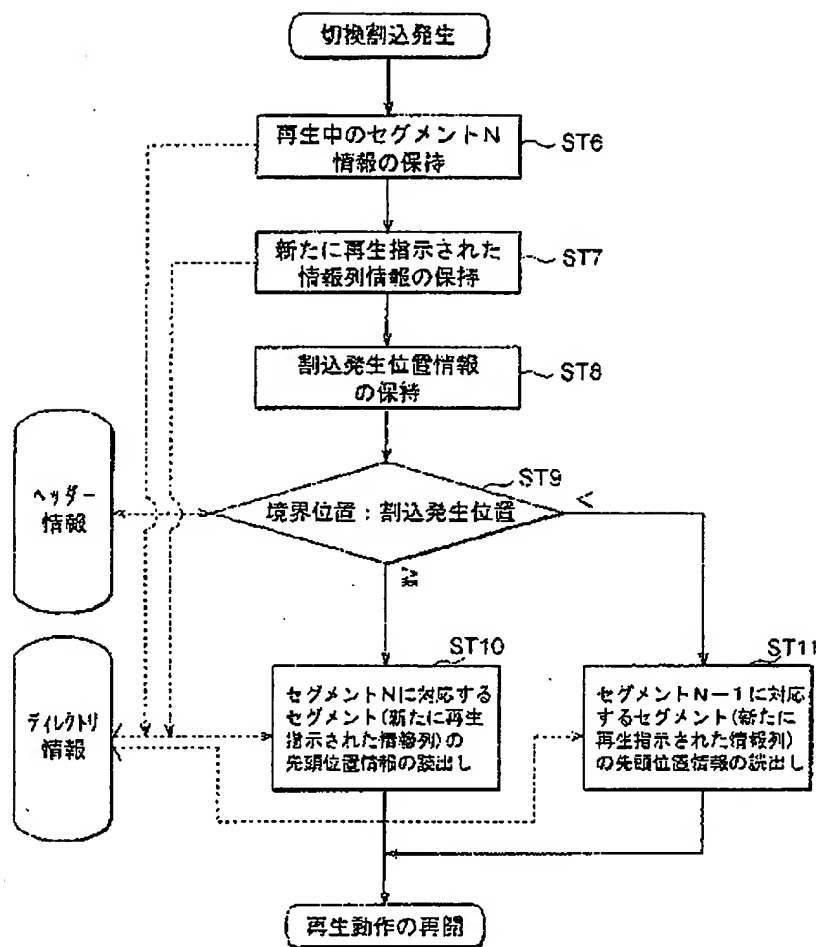
【図17】



(20)

特許2983194

【図18】





2983194

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

CLAIMS

---

(57) [Claim(s)]

[Claim 1] It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where a speech information train which was divided at a turning point of pronunciation, respectively, and which was classified into a variable-length partition for every speech information was recorded, The 2nd field where boundary location identification information in this speech information train which shows a boundary with a portion for every partition in a part for the first portion and the second half of this each partition was recorded In an interrupt-processing method under voice playback for performing processing of an interrupt request generated during speech information playback of a voice record medium which it had at least in which various playback methods are directed As opposed to an interrupt request generated during playback of a predetermined partition among said partitions this interrupt request [ whether it generated in a part for the first portion from a head location of this partition that was under playback to a boundary location, and ] Or it judges whether it generated in a second half portion from a boundary location of this partition that was under playback to a head location of degree partition. Said interrupt request judged to have generated in a part for the first portion of said partition which was under playback based on a decision result about generating timing of said interrupt request as a target partition An interrupt-processing method under voice playback said whose interrupt request judged to have generated in a second half portion of a partition which specified segmentum-antierius drawing of this partition that was under playback, or was [ said ] under playback is characterized by pinpointing the partition itself which was [ said ] under playback as a target partition.

\* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] This invention relates to the interrupt-processing method for processing the interrupt request which directs the voice record medium with which the various information which contains speech information in data medium, such as CD-ROM and DAT, at least was recorded, the playback method which reproduces the speech information beforehand recorded on such a voice record medium, and the various playback methods generated during voice playback.

[0002]

[Description of the Prior Art] From the former, the teaching materials with which speech information was recorded on record media, such as a cassette tape, for the object for self-study of language studies, such as English conversation, the object for practice of \*\*\*\*, the object for legal self-study, and the other purposes are offered variously. Here, when the teaching materials for self-study of English conversation were explained to the example, the conventional main record media are the cassette tapes (or record) with which utterance (speech information) of a series of English was recorded, and the student was using it combining this tape teaching materials and text. In addition, various level is prepared for such teaching materials from the object for the beginners' classes to the object for upper classes.

[0003] Moreover, the 1st field where the speech information train suitable for the upper person study classified into two or more partitions was recorded on the Japan patent No. 2581700, The 2nd field where the speech information train suitable for the beginners' class person study which consists of an equivalent partition corresponding to each [ these ] partition was recorded, The relation of each partition where this object for upper person study and each speech information train for beginners' class person study correspond The playback method including change playback of information record media, such as CD-ROM equipped with the 3rd field where the information shown in the record location in the record medium of each partition of these speech information train was recorded at least, and the information record medium equipped with such structure etc. is proposed.

[0004] On the other hand, the technology for realizing the repeat function under voice playback using CD for linguistic study is proposed by JP, 5-224581, A. That is, specifically, this technology is reproducing data medium by which sequential record of the speech information train of every [ to which the index number was given beforehand ] one sentence (sentence) was carried out using a predetermined

regenerative apparatus about the delay sensing function which senses the directions delay from the student under usual hearing study. For example, immediately after hearing a difficult sentence and a sentence including the pronunciation which is hard to catch, even if it is the case where the function as for which repeat playback is made to such a regenerative apparatus is prepared, repeat directions (device actuation by the student) were overdue, playback of the following sentence started, and the fault of it becoming impossible to hear the sentence which it already finished hearing was in the student. Then, although the speech information of the segment which serves as a candidate for playback by changing the index number of each [ of A B, C, and D ] segment which serves as a candidate for playback at the time one by one is reproduced with this conventional technology as shown in drawing 19, the index number (the index number which shows the speech information which carries out repeat playback) which shows a front segment to coincidence from playback initiation of the following speech information to time-amount  $\Delta t$  progress is also held by the predetermined maintenance means. If it puts in another way, this conventional technology holds independently the index number of the segment which is a candidate for playback, and the index number of the segment which should be the object of repeat playback, and they are carrying out time management so that only predetermined time  $\Delta t$  may shift the change timing of each [ these ] index number. Thus, by constituting, the jump signal (information which directs a playback starting position) corresponding to the index number currently held when the repeat demand from a student was published was generated, and the problem (directions delay) resulting from the delay of the repeat demand from a student is avoided.

[0005]

[Problem(s) to be Solved by the Invention] As mentioned above, by the conventional voice playback method, beforehand in consideration of the directions delay from an operator (for example, linguistic student) like a repeat function, delay sensing of the directions from this operator is performed so that repeat playback can be performed from the sentence which the operator meant.

[0006] However, such a delay sensing function is the technology supposing the case where a predetermined sentence is recorded regularly one by one. Like [ it follows, for example, ] a performer's conversation in a movie, and the conversation in an everyday living environment It consists of two or more sentences from which length differs. Moreover, between each sentence (each conversation person's speech information) In the playback of a series of speech information the non-voice period which occurs in random, such as a condition that voice is not reproduced, a condition that only the noise is reproduced, and a condition that only music (BGM) is played, can recognize [ playback ] long duration (for example, 1 seconds or more) existence It is difficult to realize exactly operating [ which the student who is an operator wishes in the delay sensing function by time management ]. That is, under such a condition, the case where the sentence for a repeat meant from the operator differed from the sentence by which repeat playback is actually carried out happened, as a result, by linguistic study, a student's attention was taken by device actuation and the technical problem that sufficient study effect will decrease (concentration becomes diffuse) occurred.

[0007] Then, even if it is the case where a series of speech information which consisted of two or more sentences from which it is made in order that this invention may solve the above technical problems, and length differs is reproduced, it aims at offering the voice record medium equipped with the structure which

enables playback of the speech information which an operator means exactly, its playback method, and the interrupt-processing method under voice playback.

[0008]

[Means for Solving the Problem] This invention like a performer's conversation in a movie, and conversation in an everyday living environment It consists of two or more sentences from which length differs. Moreover, between each sentence (each conversation person's speech information) Also in playback of a series of speech information in which a non-voice period which occurs in random, such as a condition that voice is not reproduced, a condition that only a noise is reproduced, and a condition that only music (BGM) is played, may exist It is related with technology of also realizing directions delay exactly in consideration of playback of speech information which an operator means.

[0009] Especially, it is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means, and one kind or two kinds or more of speech information trains which were divided at a turning point of pronunciation, respectively and which were classified into a variable-length partition for every speech information are recorded on a voice record medium which is a candidate for playback at least. In addition, in this specification, speech information is a concept and information containing a sound (voice) which can actually be caught, and a speech information train means the data itself actually recorded on record media, such as CD-ROM.

[0010] Concretely a voice record medium (the 1st embodiment concerning a voice record medium) concerning this invention It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where a speech information train which was divided, respectively at a turning point (take breath in the end of a sentence or a sentence break on \*\*, generating, or syntax) of pronunciation, and which was classified into a variable-length partition for every speech information was recorded, It has at least the 2nd field where boundary location identification information which shows a boundary with a portion in a part for the first portion and the second half of this each partition was recorded for every partition of this speech information train.

[0011] In addition, the speech information record medium concerned may be equipped with two or more fields where two or more kinds of speech information trains were recorded. As shown in the above-mentioned Japan patent No. 2581700, two or more kinds of speech information changes a delay sensing function under voice playback. Namely, when refreshable, A voice record medium (the 2nd embodiment concerning a voice record medium) concerning this invention It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where the 1st speech information train which was divided at a turning point of pronunciation, respectively, and which was classified into a variable-length partition for every speech information was recorded, And the 2nd field where the 2nd speech information train which is a speech information train which consisted of pronunciation which was equivalent to this 1st speech information train, and \*\*\*\*(ed), and was classified into a variable-length partition for every speech information equivalent to speech information of each partition of this 1st speech information train was recorded, The 3rd field where

record location identification information which shows this 1st and 2nd speech information train in a record location of each of this partition [ in / for each partition which can be switched / the record medium concerned ] that it should switch and reproduce with this voice playback means was recorded, And it reached this one and has at least the 4th field where boundary location identification information in the 2nd speech information train which shows a boundary of a portion for every partition in a part for the first portion and the second half of this each partition was recorded.

[0012] Moreover, even if a voice record medium concerning this invention is which an above-mentioned configuration, boundary location identification information, which shows a boundary of a portion a part for the first portion divided as information about a partition according to a pronunciation condition under each partition (every settlement of pronunciation) and the second half is characterized by what was recorded on a predetermined field of the voice record medium concerned.

[0013] On the other hand, a voice playback method concerning this invention specifies an object of this interrupt request using the above-mentioned boundary location identification information beforehand set as each partition to an interrupt request which directs various playback methods generated during voice playback. In addition, a repeat function, change regenerative functions which change two or more speech information stored in each field of the above-mentioned voice record medium, and are reproduced, or these combination are included in playback directions of an interrupt request at least.

[0014] When realizing an above-mentioned repeat function, by playback method (the 1st embodiment concerning a playback method) of a voice record medium concerning this invention First, it is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where a speech information train which was divided at a turning point of pronunciation, respectively, and which was classified into a variable-length partition for every speech information was recorded, A voice record medium equipped with the 2nd field where boundary location identification information which shows a boundary of a portion in a part for the first portion and the second half of this partition divided by pronunciation condition was recorded at least is prepared for every partition of this speech information train. And if an interrupt request about repeat playback from an operator is published during playback of a predetermined partition among two or more recorded partitions, generating timing of this interrupt request will be judged. from a boundary location of whether an interrupt request specifically occurred in a part for the first portion of this partition from a head location of a partition which was under playback to a boundary location, and a partition which was [ this ] under playback — since — it is judged whether an interrupt occurred in a second half portion of this partition to a head location of degree partition. Here, when it is judged that it generated in a part for the first portion of a partition which the above-mentioned interrupt request was reproducing, segmentum-antierius drawing of a partition which an interrupt request was reproducing as a target partition is specified. When it is judged that it generated on the other hand in a second half portion of a partition which the above-mentioned interrupt request was reproducing, the partition itself which an interrupt request was reproducing as a target partition is pinpointed.

[0015] In the 1st embodiment concerning a playback method, a repeat function is

realized by resuming voice playback from a head of a partition (partition pinpointed as an object partition which an operator meant) pinpointed by the above interrupt processing.

[0016] When, changing two or more speech information trains recorded on the record medium concerned on the other hand and reproducing, by playback method (the 2nd embodiment concerning a playback method) of a voice record medium concerning this invention First, it is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where the 1st speech information train which was divided at a turning point of pronunciation, respectively, and which was classified into a variable-length partition for every speech information was recorded, The 2nd field where the 2nd speech information train which is a speech information train which consisted of pronunciation which was equivalent to this 1st speech information train, and \*\*\*\* (ed), and was divided into a variable-length partition for every speech information equivalent to speech information of each partition of this 1st speech information train was recorded, The 3rd field where record location identification information which shows this 1st and 2nd speech information train in a record location of each of this partition [ in / for each partition which can be switched / the record medium concerned ] that it should switch and reproduce with this voice playback means was recorded, And a voice record medium equipped with the 4th field where boundary location identification information which shows a boundary of a portion in a part for the first portion and the second half of this each partition in this 1st and 2nd speech information train divided by pronunciation condition for every partition was recorded at least is prepared. And when an interrupt request about change playback between the 1st and 2nd speech information trains occurs during playback of a predetermined partition among these partitions, generating timing of an interrupt request is once judged. That is, it is judged whether it generated in the second-half portion of this partition from a boundary location of a partition which was [ generating / it /-in part for the first portion of this partition from head location of partition which this interrupt request was reproducing to boundary location, and this ] under playback to a head location of degree partition, and the partition which serves as the candidate for interruption based on this decision result is pinpointed. When it is judged that a partition pinpointed was generated in a part for the first portion of a partition which this interrupt request was reproducing, segmentum-antierius drawing of this partition is specified, and when it is judged that it generated in a second half portion of a partition which an interrupt request was reproducing, this partition itself is pinpointed.

[0017] In the 2nd embodiment concerning a playback method, if a partition as a referent which an operator meant by the above interrupt processing is pinpointed, an interrupt request is a partition corresponding to a partition pinpointed as a target partition, it will change from a head of a partition of a speech information train directed by the above-mentioned interrupt request among the 1st and 2nd speech information trains by resuming voice playback, and a regenerative function will be realized.

[0018] In addition, even if it changes as mentioned above and is the regeneration method of a refreshable voice record medium, it is possible to have various functions, such as above-mentioned repeat playback, further.

[0019] Moreover, a voice record medium concerning this invention contains a refreshable information record medium with a personal computer equipped with a



portable CD player, a CD-ROM driver, etc. Therefore, corresponding to playback timing of each speech information, you may also include refreshable image information, alphabetic information, etc. in order to be reproduced by such various devices.

[0020]

[Embodiment of the Invention] Hereafter, one example of this invention is explained using drawing 1 - drawing 18. In addition, the same sign is given to a same-among drawing portion, and explanation is omitted.

[0021] (The 1st embodiment concerning a voice record medium) The fundamental structure of the speech information train which should be first recorded on the voice record medium concerning this invention is notionally explained using drawing 1.

[0022] The information recorded with the voice record medium concerning this invention Like a performer's conversation in a movie, and the conversation in an everyday living environment It consists of two or more sentences (sentence) from which length differs. Moreover, between each sentence (each conversation person's speech information) They are a series of speech information trains in which the non-voice period which occurs in random, such as a condition that voice is not reproduced, a condition that only the noise is reproduced, and a condition that only music (BGM) is played, may exist. Therefore, the 1st embodiment (basic configuration) concerning this voice record medium It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. It has at least the 1st field where the speech information train which was divided at the turning point of pronunciation, respectively, and which was classified into the variable-length partition (henceforth a segment) for every speech information was recorded.

[0023] Since one sentence is generally uttered in general in about 3 seconds in a native speaker's English conversation, it is appropriate to constitute the variable-length segment which constitutes a speech information train, respectively from setting up the turning point of the pronunciation which determines the segment which constitutes the speech information train which should be recorded between each sentence, as shown in drawing 1 (a), (b), or (d). In addition, in an all busy sentence, as shown in drawing 1 (c), an extremely short sentence is contained, and this sentence also constitutes one segment. It is appropriate to constitute from two continuous segments for a sentence as shown in drawing 1 (e) on the other hand, in the case of an extremely long sentence, since a conjunction, relative, etc. front serves as a turning point of pronunciation as shown in drawing 1 (e). Therefore, the segment of the speech information train which should be recorded means that it is the record unit of the speech information divided based on the break on utterance (breathing location), or a certain break on language (on syntax).

[0024] Moreover, the voice record medium concerned is further equipped with the 2nd field where the boundary location identification information which shows the boundary of a portion in a part for the first portion and the second half of each segment divided at the turning point of pronunciation as mentioned above that the interrupt request which may be generated during playback of the above special speech information should be processed appropriately was recorded. This boundary location identification information is concretely recorded, for example on the header unit of the directory field of the voice record medium concerned, or each

segment.

[0025] The divided boundary which is set up for every segment As shown in drawing 1 (a), (b), or (d), it is set as the first breaks S1 and S2 with a settlement of a certain pronunciation, and S4 (). That is, S1, S2, S3, and S4 are prepared according to a pronunciation condition for every partition, and, thereby, each partition is divided into a portion a part for the first portion, and the second half. In addition, as shown in drawing 1 (c), when a sentence is extremely short, you may set up so that it may be in agreement with the termination location of the segment divided in the boundary location S3. Moreover, as shown in drawing 1 (e), also when one sentence is divided into two or more segments, the boundary locations S5 and S6 which show the boundary divided into a portion a part for the first portion and the second half in which it was judged as the settlement of pronunciation about each segment are set up, respectively.

[0026] Especially voice playback of the voice record medium concerning this 1st embodiment is performed in order for every recorded segment. However, when the interrupt request from operators, such as repeat directions, occurs during the starting voice playback, the generating timing of this interrupt request is once judged, and the segment set as the object of this interrupt request based on this decision result is specified. It is judged whether it specifically generated in the second half portion to the head location of degree segment which should be reproduced next from the boundary location of the segment which was [ generating / it /-in part for the first portion from head location of segment which this interrupt request was reproducing to boundary location set up as mentioned above, and this ] under playback. Here, when it is judged that it generated in a part for the first portion of the segment which the above-mentioned interrupt request was reproducing, the segment in front of the segment which the interrupt request was reproducing as a target segment (segment which voice playback already ended) is specified. When it is judged that it generated in the portion on the other hand in the second half of a segment in which the above-mentioned interrupt request was being reproduced, the segment itself which the interrupt request was reproducing as a target segment is specified.

[0027] This playback method realizes a repeat function by resuming voice playback from the head of the segment (segment specified as an operator's referent) specified in this way.

[0028] (The 2nd embodiment concerning a voice record medium) Next, the 2nd embodiment of the voice record medium concerning this invention is characterized by recording at least two kinds of speech information trains. That is, the 1st speech information train consists of English speech information about which a native speaker speaks with a natural speed, and this speech information train is divided into two or more variable-length segments at the turning point (take breath in the end of a sentence, or a sentence break on \*\*, generating, or syntax) of pronunciation as mentioned above. Although the 2nd speech information train is a semantic content equivalent to the contents of the above-mentioned 1st speech information train, it is another speech information, for example, it is speech information of English of the speed carried out slowly which divides and speaks a word. In addition, this 2nd speech information train also consists of two or more variable-length segments.

[0029] An important thing is that the semantic content corresponds for every segment mutually here, although the above-mentioned 1st and 2nd speech information train is classified into two or more variable-length segments, respectively. For

example, when the  $t$ -th segment ( drawing 1 (a) the 621st) of the 1st speech information train is "It's not much of a problem." about which a native speaker speaks, the  $t$ -th segment of the 2nd speech information train becomes "It is not much of a problem." which divides and speaks each word. However, it is shown that consisting of another speech information which is the 2nd speech information train and corresponding contents is that from which utterance differs in semantics with the same language top.

[0030] Furthermore, the voice record medium concerning this 2nd embodiment is characterized by recording record location identification information on the predetermined field. This record location identification information is information which shows on which location of the voice record medium concerned each segment in the 1st and 2nd speech information train is recorded at least. therefore, the  $t$ -th segment of for example, the 1st speech information train — it can be recognized by this record location identification information on which location of data medium "segment "It is not much of a problem. of the 2nd speech information train corresponding to It's not much of a problem."" is recorded.

[0031] Consequently, the 1st and 2nd speech information train and record location identification information are not recorded independently mutually, but it is recorded with fixed relation, and each speech information train is together put organically considering the segment as an unit. That is, the 1st and 2nd speech information train is making the pair mutually, and it is record location identification information to relate these for every segment. In addition, in this embodiment, record record location identification information is recorded on the directory field of the voice record medium concerned, and includes the information about the head location of each segment at least. Moreover, in this embodiment, boundary location identification information is recorded on a part for the header unit of each segment.

[0032] Although voice playback is performed in order for every recorded segment also by the playback method of the voice record medium (the 2nd embodiment) equipped with the above structures, by this playback method, it is especially characterized by the playback change (or playback change in the 1st speech information train from the 2nd speech information train) in the 2nd speech information train recorded on the voice record medium concerned from the 1st speech information train being possible. In addition, this playback transfer operation is performed considering a segment as an unit. For example, if playback directions of the 2nd speech information train are inputted when the  $t$ -th segment of the 1st speech information train is reproduced (generating of an interrupt request), the  $t$ -th segment to which the 2nd speech information train corresponds based on record location identification information will be read, and voice playback of the corresponding segment will be performed. Moreover, it is carried out in each segment unit like the playback transfer operation which also mentioned above the playback change in the 1st speech information train from the 2nd speech information train conversely.

[0033] In addition, in above-mentioned transfer operation, specification of the segment used as the candidate for a change is performed by judging that it was explained by the playback method of the voice record medium concerning the 1st embodiment which had interruption timing mentioned above. Moreover, by this playback method, various deformation of repeat playback besides above-mentioned playback transfer operation etc. is possible. The so-called return command is one of the typical thing. That is, when it returns after interrupting playback by the

stop instruction under playback temporarily, and a command is inputted, playback of the speech information only whose ordered amount suited to an operator's hope more by returning the read-out location of speech information is performed.

[0034] (The 3rd embodiment of a voice record medium) the 3rd operative condition of the voice record medium further applied to this invention — it attaches like and explains. The voice record medium concerning this 3rd embodiment differs from the voice record medium which requires for the 2nd embodiment of the above the point that the 3rd speech information train is further recorded on the voice record medium concerned besides the 1st and 2nd speech information train.

[0035] An important thing is that the above-mentioned 3rd speech information train is classified into the segment group which made the bundle 1 or two or more variable-length segments of the 1st and 2nd speech information train here. If it puts in another way, one segment group of this 3rd speech information train includes 1 or two or more segments of the 1st and 2nd speech information train, therefore one segment group of the 3rd speech information train has become 1 or two or more segments of the 1st and 2nd speech information train, and a pair.

Especially this configuration assumes the case where one sentence is classified into two or more segments, as shown in drawing 1 (e).

[0036] Moreover, in the voice record medium of this 3rd embodiment, the information which also shows the record location of the contents of the above-mentioned 3rd speech information train for every segment group is also included in the record location identification information recorded on the predetermined field. Therefore, the 1st, 2nd, and 3rd speech information train and record location identification information are mutually recorded on data medium with fixed relation, and each speech information train is together put organically considering the segment or the segment group as an unit. In addition, also in this 3rd embodiment, record location identification information is recorded on the directory field of the voice record medium concerned, and includes the information about the head location of the segment in each speech information train. Moreover, also in this embodiment, boundary location identification information is recorded on the header unit of each segment.

[0037] the 2nd operative condition by which the playback method of the voice record medium (the 3rd embodiment) equipped with the above structures was fundamentally mentioned above — the point of performing playback transfer operation between the others and this 1st and 2nd speech information trains and the 3rd speech information trains although it is the same as the case like — the 2nd operative condition — it differs from the case like. [ change / between the 1st and 2nd speech information trains / playback ]

[0038] For example, when a native speaker's "It's not much of a problem." is not able to be caught during playback of the 1st speech information train, "It is not much of a problem." which divides and speaks a word slowly can be heard by switching playback to the 2nd speech information train from the 1st speech information train under playback. And what is necessary is just to switch playback to the 3rd speech information train from the speech information train under playback further to know the semantics and the syntax of this Japanese. Of course, in this playback method, the thing which was explained by the playback method of the voice record medium concerning the 2nd embodiment of the above and to which it can apply so that it may return and can use combining a command or stop instruction cannot be overemphasized. Moreover, also in this playback method, change playback and repeat playback are possible, and specification of the target segment

[ interrupt requests / these ] is performed like the case of the 1st embodiment mentioned above.

[0039] (The 4th embodiment of a voice record medium) the 4th embodiment of the voice record medium concerning this invention — fundamental — the 3rd operative condition of the above — although it is the same as that of the case like, it mainly differs in that the alphabetic information train besides the 1st and 2nd speech information train is recorded. This alphabetic information train is equivalent to the alphabetic information corresponding to English (voice) which is equivalent to the alphabetic information of the contents corresponding to the 1st or 2nd speech information train, for example, a native speaker speaks.

[0040] This alphabetic information train is also classified into each segment of the 1st and 2nd speech information train, and a corresponding segment. moreover, this 4th operative condition — also in the voice record medium applied like, the information about those head locations is included in record location identification information for every segment of each speech information train in the record location of this alphabetic information train, and it is recorded on the directory field of the voice record medium concerned. Moreover, boundary location identification information is recorded on the header unit of each segment.

Therefore, the 1st and 2nd speech information train and an alphabetic information train will correspond per segment, respectively. in addition, this 4th operative condition — the voice record medium applied like — setting — the 3rd above-mentioned operative condition — when adding the 3rd speech information train which can be set like as recording information, 1 or two or more segments of the 1st and 2nd speech information train and an alphabetic information train will correspond also to one segment group of the 3rd speech information train. Furthermore, also in this 4th embodiment, the head location of each segment is included in the above-mentioned record location identification information, and it is recorded on the directory field of the voice record medium concerned. Moreover, boundary location identification information is recorded on the header unit of each segment.

[0041] the playback method of the voice record medium (the 4th embodiment) equipped with the above structures — fundamental — the 2nd operative condition of the above — although it is the same as that of the case like, it differs in that a display indication of the alphabetic information train is given during playback of the 1st or 2nd speech information train.

[0042] For example, when segment "It's not much of a problem." of the 1st speech information train is reproduced, a display indication of "It's not much of a problem." or the "It is not much of a problem." is given at a predetermined display. In addition, it is not necessary to synchronize with the speech information train and time amount target under playback completely about this display, and an alphabetic character may be displayed little by little behind time, or may be displayed little by little first. Moreover, also in this playback method, change playback and repeat playback are possible, and specification of the target segment [ interrupt requests / these ] is performed like the playback method of the voice record medium concerning the 1st embodiment mentioned above.

[0043] Next, the concrete structure of the voice record medium concerning this invention is explained to details below using drawing 2 - drawing 5.

[0044] Drawing 2 is drawing for explaining each speech information trains A, B, and C and the contents of record of those when applying the 3rd above-mentioned embodiment to English-conversation self-study as an example of the voice record medium concerning this this invention. In this drawing, the speech information

train A is an English information train (the 1st speech information train) about which a native speaker speaks, and consists of two or more segments 621 and 622. It is the information train (the 2nd speech information train) of English which divides a word and is spoken slowly, and, as for the speech information train B, this also consists of two or more segments which consist of the previous alphabetic word or previous phrase corresponding to segments 621 and 622 of the speech information train A. Moreover, the segment group which the speech information train C is an information train (the 3rd speech information train) which explains Japanese, and is contained in this speech information train C supports each segments 621 and 622 of each speech information trains A and B, respectively.

[0045] Moreover, drawing 3 is a table for explaining the time amount per one segment and the relation of capacity to the mode shown in drawing 2. In this table, the capacity of 6 K bytes is supported for 1 second. for example, in the segment 621 of the speech information train A The capacity for 0.2 seconds 1.2KB (kilobyte), [ the utterance time amount of "It's" ] The capacity for 0.1 seconds 0.6KB (kilobyte), [ the utterance time amount of "not" ] The capacity for 0.4 seconds 2.4KB (kilobyte), [ the utterance time amount of "much of a" ] And the utterance time amount of "problem" is 0.3 seconds, the capacity is 1.8KB (kilobyte), the utterance time amount of the segment 621 whole sets, and the capacity is set to 12KB (kilobyte) for 2.0 seconds. In addition, the numeric character of the parenthesis in a table is the value which showed the boundary location by playback time amount. That is, in the case of this speech information train A, since a boundary location is set up after "not" as shown in drawing 1 (a), it is shown that the boundary which divides the speech information train A into a portion a part for the first portion and the second half is set as the location 0.3 seconds after playback initiation of a segment 621. Moreover, although other speech information trains B and C are shown similarly, in the embodiment of this drawing 3, boundary location identification information is not set to the speech information train C.

[0046] Furthermore, drawing 4 is a table for explaining the contents of record of the directory field in the mode shown in drawing 2 and drawing 3. In this table, a directory field consists of  $9 \times 3 = 27$  byte (B) per one segment. The speech information trains A, B, and C are equivalent to the speech information trains A, B, and C of drawing 2; respectively. Moreover, 1 byte of C shows an attribute and it means that  $C = 0$  is the speech information train A, and  $C = 64$  is the speech information train B. moreover,  $C = \text{---}$  it means that 128 and 129 are the speech information trains C, and shows especially that it is the same candidate for description as a front segment at the time of "10000001" (belonging to the same segment group set as the description object of the speech information train C is shown, for example, the case of the segments 801 and 802 of drawing 1 (e) corresponds)., the time (8 bits (bit)), i.e., a bit expression, of  $C = 129$

[0047]



\* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

TECHNICAL FIELD

---

[The technical field to which invention belongs] This invention relates to the interrupt-processing method for processing the interrupt request which directs the voice record medium with which the various information which contains speech information in data medium, such as CD-ROM and DAT, at least was recorded, the playback method which reproduces the speech information beforehand recorded on such a voice record medium, and the various playback methods generated during voice playback.

---

[Translation done.]

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

## PRIOR ART

---

[Description of the Prior Art] From the former, the teaching materials with which speech information was recorded on record media, such as a cassette tape, for the object for self-study of language studies, such as English conversation, the object for practice of \*\*\*\*, the object for legal self-study, and the other purposes are offered variously. Here, when the teaching materials for self-study of English conversation were explained to the example, the conventional main record media are the cassette tapes (or record) with which utterance (speech information) of a series of English was recorded, and the student was using it combining this tape teaching materials and text. In addition, various level is prepared for such teaching materials from the object for the beginners' classes to the object for upper classes.

[0003] Moreover, the 1st field where the speech information train suitable for the upper person study classified into two or more partitions was recorded on the Japan patent No. 2581700. The 2nd field where the speech information train suitable for the beginners' class person study which consists of an equivalent partition corresponding to each [ these ] partition was recorded. The relation of each partition where this object for upper person study and each speech information train for beginners' class person study correspond. The playback method including change playback of information record media, such as CD-ROM equipped with the 3rd field where the information shown in the record location in the record medium of each partition of these speech information train was recorded at least, and the information record medium equipped with such structure etc. is proposed.

[0004] On the other hand, the technology for realizing the repeat function under voice playback using CD for linguistic study is proposed by JP,5-224581,A. That is, specifically, this technology is reproducing data medium by which sequential record of the speech information train of every [ to which the index number was given beforehand ] one sentence (sentence) was carried out using a predetermined regenerative apparatus about the delay sensing function which senses the directions delay from the student under usual hearing study. For example, immediately after hearing a difficult sentence and a sentence including the pronunciation which is hard to catch, even if it is the case where the function as for which repeat playback is made to such a regenerative apparatus is prepared, repeat directions (device actuation by the student) were overdue, playback of the following sentence started, and the fault of it becoming impossible to hear the sentence which it already finished hearing was in the student. Then, although the speech information of the segment which serves as a candidate for playback by changing the index number of each [ of A B, C, and D ] segment which serves as a candidate for playback at the time one by one is reproduced with this conventional technology as shown in drawing 19, the index number (the index number which shows the speech information which carries out repeat playback) which shows a front segment to coincidence from playback initiation of the following speech information to time-amount deltat progress is also held by the predetermined maintenance means. If it puts in another way, this conventional technology holds independently the index number of the segment which is a candidate for playback, and the index number of the segment which should be the object of repeat playback, and they are carrying out time management so that only predetermined time deltat may shift the change timing of each [ these ] index number. Thus, by constituting, the jump signal

(information which directs a playback starting position) corresponding to the index number currently held when the repeat demand from a student was published was generated, and the problem (directions delay) resulting from the delay of the repeat demand from a student is avoided.

---

[Translation done.]

\* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

EFFECT OF THE INVENTION

---

[Effect of the Invention] As information about the variable-length segment which constitutes this speech information train from this invention with various speech information as mentioned above The voice record medium with which the boundary location identification information which shows a boundary with a portion a part for the first portion divided according to the pronunciation condition of each of this segment in a record medium and the second half was recorded at least is prepared. Boundary location before of the segment which the timing of a predetermined interrupt request was reproducing, or the back is judged, and the object segment of an interrupt request is specified based on this decision result. Therefore, it is effective in the flexible playback actuation as which an operator's directions delay was also considered being realizable.

---

[Translation done.]

\* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

TECHNICAL PROBLEM

---

[Problem(s) to be Solved by the Invention] As mentioned above, by the conventional voice playback method, beforehand in consideration of the directions delay from an operator (for example, linguistic student) like a repeat function, delay sensing of the directions from this operator is performed so that repeat playback can be performed from the sentence which the operator meant.

[0006] However, such a delay sensing function is the technology supposing the case where a predetermined sentence is recorded regularly one by one. Like [ it follows, for example, ] a performer's conversation in a movie, and the conversation in an everyday living environment It consists of two or more sentences from which length differs. Moreover, between each sentence (each conversation person's speech information) In the playback of a series of speech information the non-voice period which occurs in random, such as a condition that voice is not reproduced, a condition that only the noise is reproduced, and a condition that only music (BGM) is played, can recognize [ playback ] long duration (for example, 1 seconds or more) existence It is difficult to realize exactly operating [ which the student who is an operator wishes in the delay sensing function by time management ]. That is, under such a condition, the case where the sentence for a repeat meant from the operator differed from the sentence by which repeat playback is actually carried out happened, as a result, by linguistic study, a student's attention was taken by device actuation and the technical problem that sufficient study effect will decrease (concentration becomes diffuse) occurred.

[0007] Then, even if it is the case where a series of speech information which consisted of two or more sentences from which it is made in order that this invention may solve the above technical problems, and length differs is reproduced, it aims at offering the voice record medium equipped with the structure which enables playback of the speech information which an operator means exactly, its playback method, and the interrupt-processing method under voice playback.

---

[Translation done.]

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

MEANS

---

[Means for Solving the Problem] This invention like a performer's conversation in a movie, and conversation in an everyday living environment It consists of two or more sentences from which length differs. Moreover, between each sentence (each conversation person's speech information) Also in playback of a series of speech information in which a non-voice period which occurs in random, such as a condition that voice is not reproduced, a condition that only a noise is reproduced, and a condition that only music (BGM) is played, may exist It is related with technology of also realizing directions delay exactly in consideration of playback of speech information which an operator means.

[0009] Especially, it is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means, and one kind or two kinds or more of speech information trains which were divided at a turning point of pronunciation, respectively and which were classified into a variable-length partition for every speech information are recorded on a voice record medium which is a candidate for playback at least. In addition, in this specification, speech information is a concept and information containing a sound (voice) which can actually be caught, and a speech information train means the data itself actually recorded on record media, such as CD-ROM.

[0010] Concretely a voice record medium (the 1st embodiment concerning a voice record medium) concerning this invention It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where a speech information train which was divided, respectively at a turning point (take breath in the end of a sentence or a sentence break on \*\*, generating, or syntax) of pronunciation, and which was classified into a variable-length partition for every speech information was recorded, It has at least the 2nd field where boundary location identification information which shows a boundary with a portion in a part for the first portion and the second half of this each partition was recorded for every partition of this speech information train.

[0011] In addition, the speech information record medium concerned may be equipped with two or more fields where two or more kinds of speech information trains were recorded. As shown in the above-mentioned Japan patent No. 2581700, two or more kinds of speech information changes a delay sensing function under voice playback. Namely, when refreshable, A voice record medium (the 2nd embodiment concerning a voice record medium) concerning this invention It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where the 1st speech information train which was divided at a turning point of pronunciation, respectively, and which was classified into a variable-length partition for every speech information was recorded, And the 2nd field where the 2nd speech information train which is a speech information train which consisted of pronunciation which was equivalent to this 1st speech information train, and \*\*\*\*(ed), and was classified into a variable-length partition for every speech information equivalent to speech information of each partition of this 1st speech information train was recorded, The 3rd field where record location identification information which shows this 1st and 2nd speech information train in a record location of

each of this partition [ in / for each partition which can be switched / the record medium concerned ] that it should switch and reproduce with this voice playback means was recorded, And it reached this one and has at least the 4th field where boundary location identification information in the 2nd speech information train which shows a boundary of a portion for every partition in a part for the first portion and the second half of this each partition was recorded.

[0012] Moreover, even if a voice record medium concerning this invention is which an above-mentioned configuration, boundary location identification information which shows a boundary of a portion a part for the first portion divided as information about a partition according to a pronunciation condition under each partition (every settlement of pronunciation) and the second half is characterized by what was recorded on a predetermined field of the voice record medium concerned.

[0013] On the other hand, a voice playback method concerning this invention specifies an object of this interrupt request using the above-mentioned boundary location identification information beforehand set as each partition to an interrupt request which directs various playback methods generated during voice playback. In addition, a repeat function, change regenerative functions which change two or more speech information stored in each field of the above-mentioned voice record medium, and are reproduced, or these combination are included in playback directions of an interrupt request at least.

[0014] When realizing an above-mentioned repeat function, by playback method (the 1st embodiment concerning a playback method) of a voice record medium concerning this invention First, it is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where a speech information train which was divided at a turning point of pronunciation, respectively, and which was classified into a variable-length partition for every speech information was recorded, A voice record medium equipped with the 2nd field where boundary location identification information which shows a boundary of a portion in a part for the first portion and the second half of partition divided by pronunciation condition was recorded at least is prepared for every partition of this speech information train. And if an interrupt request about repeat playback from an operator is published during playback of a predetermined partition among two or more recorded partitions, generating timing of this interrupt request will be judged. from a boundary location of whether an interrupt request specifically occurred in a part for the first portion of this partition from a head location of a partition which was under playback to a boundary location, and a partition which was [ this ] under playback -- since -- it is judged whether an interrupt occurred in a second half portion of this partition to a head location of degree partition. Here, when it is judged that it generated in a part for the first portion of a partition which the above-mentioned interrupt request was reproducing, segmentum-antierius drawing of a partition which an interrupt request was reproducing as a target partition is specified. When it is judged that it generated on the other hand in a second half portion of a partition which the above-mentioned interrupt request was reproducing, the partition itself which an interrupt request was reproducing as a target partition is pinpointed.

[0015] In the 1st embodiment concerning a playback method, a repeat function is realized by resuming voice playback from a head of a partition (partition pinpointed as an object partition which an operator meant) pinpointed by the above interrupt processing.

[0016] When, changing two or more speech information trains recorded on the record medium concerned on the other hand and reproducing, by playback method (the 2nd embodiment concerning a playback method) of a voice record medium concerning this invention First, it is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The 1st field where the 1st speech information train which was divided at a turning point of pronunciation, respectively, and which was classified into a variable-length partition for every speech information was recorded, The 2nd field where the 2nd speech information train which is a speech information train which consisted of pronunciation which was equivalent to this 1st speech information train, and \*\*\*\*(ed), and was divided into a variable-length partition for every speech information equivalent to speech information of each partition of this 1st speech information train was recorded, The 3rd field where record location

identification information which shows this 1st and 2nd speech information train in a record location of each of this partition [ in / for each partition which can be switched / the record medium concerned ] that it should switch and reproduce with this voice playback means was recorded, And a voice record medium equipped with the 4th field where boundary location identification information which shows a boundary of a portion in a part for the first portion and the second half of this each partition in this 1st and 2nd speech information train divided by pronunciation condition for every partition was recorded at least is prepared. And when an interrupt request about change playback between the 1st and 2nd speech information trains occurs during playback of a predetermined partition among these partitions, generating timing of an interrupt request is once judged. That is, it is judged whether it generated in the second-half portion of this partition from a boundary location of a partition which was [ generating / it / - in part for the first portion of this partition from head location of partition which this interrupt request was reproducing to boundary location, and this ] under playback to a head location of degree partition, and the partition which serves as the candidate for interruption based on this decision result is pinpointed. When it is judged that a partition pinpointed was generated in a part for the first portion of a partition which this interrupt request was reproducing, segmentum-antierius drawing of this partition is specified, and when it is judged that it generated in a second half portion of a partition which an interrupt request was reproducing, this partition itself is pinpointed.

[0017] In the 2nd embodiment concerning a playback method, if a partition as a referent which an operator meant by the above interrupt processing is pinpointed, an interrupt request is a partition corresponding to a partition pinpointed as a target partition, it will change from a head of a partition of a speech information train directed by the above-mentioned interrupt request among the 1st and 2nd speech information trains by resuming voice playback, and a regenerative function will be realized.

[0018] In addition, even if it changes as mentioned above and is the regeneration method of a refreshable voice record medium, it is possible to have various functions, such as above-mentioned repeat playback, further.

[0019] Moreover, a voice record medium concerning this invention contains a refreshable information record medium with a personal computer equipped with a portable CD player, a CD-ROM driver, etc. Therefore, corresponding to playback timing of each speech information, you may also include refreshable image information, alphabetic information, etc. in order to be reproduced by such various devices.

[0020]

[Embodiment of the Invention] Hereafter, one example of this invention is explained using drawing 1 - drawing 18 . In addition, the same sign is given to a same-among drawing portion, and explanation is omitted.

[0021] (The 1st embodiment concerning a voice record medium) The fundamental structure of the speech information train which should be first recorded on the voice record medium concerning this invention is notionally explained using drawing 1 .

[0022] The information recorded with the voice record medium concerning this invention Like a performer's conversation in a movie, and the conversation in an everyday living environment It consists of two or more sentences (sentence) from which length differs. Moreover, between each sentence (each conversation person's speech information) They are a series of speech information trains in which the non-voice period which occurs in random, such as a condition that voice is not reproduced, a condition that only the noise is reproduced, and a condition that only music (BGM) is played, may exist. Therefore, the 1st embodiment (basic configuration) concerning this voice record medium It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. It has at least the 1st field where the speech information train which was divided at the turning point of pronunciation, respectively, and which was classified into the variable-length partition (henceforth a segment) for every speech information was recorded.

[0023] Since one sentence is generally uttered in general in about 3 seconds in a native speaker's English conversation, it is appropriate to constitute the variable-length segment which constitutes a speech



information train, respectively from setting up the turning point of the pronunciation which determines the segment which constitutes the speech information train which should be recorded between each sentence, as shown in drawing 1 (a), (b), or (d). In addition, in an all busy sentence, as shown in drawing 1 (c), an extremely short sentence is contained, and this sentence also constitutes one segment. It is appropriate to constitute from two continuous segments for a sentence as shown in drawing 1 (e) on the other hand, in the case of an extremely long sentence, since a conjunction, relative, etc. front serves as a turning point of pronunciation as shown in drawing 1 (e). Therefore, the segment of the speech information train which should be recorded means that it is the record unit of the speech information divided based on the break on utterance (breathing location), or a certain break on language (on syntax). [0024] Moreover, the voice record medium concerned is further equipped with the 2nd field where the boundary location identification information which shows the boundary of a portion in a part for the first portion and the second half of each segment divided at the turning point of pronunciation as mentioned above that the interrupt request which may be generated during playback of the above special speech information should be processed appropriately was recorded. This boundary location identification information is concretely recorded, for example on the header unit of the directory field of the voice record medium concerned, or each segment.

[0025] The divided boundary which is set up for every segment As shown in drawing 1 (a), (b), or (d), it is set as the first breaks S1 and S2 with a settlement of a certain pronunciation, and S4 (). That is, S1, S2, S3, and S4 are prepared according to a pronunciation condition for every partition, and, thereby, each partition is divided into a portion a part for the first portion, and the second half. In addition, as shown in drawing 1 (c), when a sentence is extremely short, you may set up so that it may be in agreement with the termination location of the segment divided in the boundary location S3. Moreover, as shown in drawing 1 (e), also when one sentence is divided into two or more segments, the boundary locations S5 and S6 which show the boundary divided into a portion a part for the first portion and the second half in which it was judged as the settlement of pronunciation about each segment are set up, respectively.

[0026] Especially voice playback of the voice record medium concerning this 1st embodiment is performed in order for every recorded segment. However, when the interrupt request from operators, such as repeat directions, occurs during the starting voice playback, the generating timing of this interrupt request is once judged, and the segment set as the object of this interrupt request based on this decision result is specified. It is judged whether it specifically generated in the second half portion to the head location of degree segment which should be reproduced next from the boundary location of the segment which was [ generating / it /-in part for the first portion from head location of segment which this interrupt request was reproducing to boundary location set up as mentioned above, and this ] under playback. Here, when it is judged that it generated in a part for the first portion of the segment which the above-mentioned interrupt request was reproducing, the segment in front of the segment which the interrupt request was reproducing as a target segment (segment which voice playback already ended) is specified. When it is judged that it generated in the portion on the other hand in the second half of a segment in which the above-mentioned interrupt request was being reproduced, the segment itself which the interrupt request was reproducing as a target segment is specified.

[0027] This playback method realizes a repeat function by resuming voice playback from the head of the segment (segment specified as an operator's referent) specified in this way.

[0028] (The 2nd embodiment concerning a voice record medium) Next, the 2nd embodiment of the voice record medium concerning this invention is characterized by recording at least two kinds of speech information trains. That is, the 1st speech information train consists of English speech information about which a native speaker speaks with a natural speed, and this speech information train is divided into two or more variable-length segments at the turning point (take breath in the end of a sentence, or a sentence break on \*\*, generating, or syntax) of pronunciation as mentioned above. Although the 2nd speech information train is a semantic content equivalent to the contents of the above-mentioned 1st speech information train, it is another speech information, for example, it is speech information of English of the speed carried out slowly which divides and speaks a word. In addition, this

2nd speech information train also consists of two or more variable-length segments.

[0029] An important thing is that the semantic content corresponds for every segment mutually here, although the above-mentioned 1st and 2nd speech information train is classified into two or more variable-length segments, respectively. For example, when the t-th segment ( drawing 1 (a) the 621st) of the 1st speech information train is "It's not much of a problem." about which a native speaker speaks, the t-th segment of the 2nd speech information train becomes "It is not much of a problem." which divides and speaks each word. However, it is shown that consisting of another speech information which is the 2nd speech information train and corresponding contents is that from which utterance differs in semantics with the same language top.

[0030] Furthermore, the voice record medium concerning this 2nd embodiment is characterized by recording record location identification information on the predetermined field. This record location identification information is information which shows on which location of the voice record medium concerned each segment in the 1st and 2nd speech information train is recorded at least. therefore, the t-th segment of for example, the 1st speech information train -- it can be recognized by this record location identification information on which location of data medium "segment "It is not much of a problem. of the 2nd speech information train corresponding to It's not much of a problem."" is recorded.

[0031] Consequently, the 1st and 2nd speech information train and record location identification information are not recorded independently mutually, but it is recorded with fixed relation, and each speech information train is together put organically considering the segment as an unit. That is, the 1st and 2nd speech information train is making the pair mutually, and it is record location identification information to relate these for every segment. In addition, in this embodiment, record record location identification information is recorded on the directory field of the voice record medium concerned, and includes the information about the head location of each segment at least. Moreover, in this embodiment, boundary location identification information is recorded on a part for the header unit of each segment.

[0032] Although voice playback is performed in order for every recorded segment also by the playback method of the voice record medium (the 2nd embodiment) equipped with the above structures, by this playback method, it is especially characterized by the playback change (or playback change in the 1st speech information train from the 2nd speech information train) in the 2nd speech information train recorded on the voice record medium concerned from the 1st speech information train being possible. In addition, this playback transfer operation is performed considering a segment as an unit. For example, if playback directions of the 2nd speech information train are inputted when the t-th segment of the 1st speech information train is reproduced (generating of an interrupt request), the t-th segment to which the 2nd speech information train corresponds based on record location identification information will be read, and voice playback of the corresponding segment will be performed. Moreover, it is carried out in each segment unit like the playback transfer operation which also mentioned above the playback change in the 1st speech information train from the 2nd speech information train conversely.

[0033] In addition, in above-mentioned transfer operation, specification of the segment used as the candidate for a change is performed by judging that it was explained by the playback method of the voice record medium concerning the 1st embodiment which had interruption timing mentioned above. Moreover, by this playback method, various deformation of repeat playback besides above-mentioned playback transfer operation etc. is possible. The so-called return command is one of the typical thing. That is, when it returns after interrupting playback by the stop instruction under playback temporarily, and a command is inputted, playback of the speech information only whose ordered amount suited to an operator's hope more by returning the read-out location of speech information is performed.

[0034] (The 3rd embodiment of a voice record medium) the 3rd operative condition of the voice record medium further applied to this invention -- it attaches like and explains. The voice record medium concerning this 3rd embodiment differs from the voice record medium which requires for the 2nd embodiment of the above the point that the 3rd speech information train is further recorded on the voice record medium concerned besides the 1st and 2nd speech information train.

[0035] An important thing is that the above-mentioned 3rd speech information train is classified into the

segment group which made the bundle 1 or two or more variable-length segments of the 1st and 2nd speech information train here. If it puts in another way, one segment group of this 3rd speech information train includes 1 or two or more segments of the 1st and 2nd speech information train, therefore one segment group of the 3rd speech information train has become 1 or two or more segments of the 1st and 2nd speech information train, and a pair. Especially this configuration assumes the case where one sentence is classified into two or more segments, as shown in drawing 1 (e).

[0036] Moreover, in the voice record medium of this 3rd embodiment, the information which also shows the record location of the contents of the above-mentioned 3rd speech information train for every segment group is also included in the record location identification information recorded on the predetermined field. Therefore, the 1st, 2nd, and 3rd speech information train and record location identification information are mutually recorded on data medium with fixed relation, and each speech information train is together put organically considering the segment or the segment group as an unit. In addition, also in this 3rd embodiment, record location identification information is recorded on the directory field of the voice record medium concerned, and includes the information about the head location of the segment in each speech information train. Moreover, also in this embodiment, boundary location identification information is recorded on the header unit of each segment.

[0037] the 2nd operative condition by which the playback method of the voice record medium (the 3rd embodiment) equipped with the above structures was fundamentally mentioned above -- the point of performing playback transfer operation between the others and this 1st and 2nd speech information trains and the 3rd speech information trains although it is the same as the case like -- the 2nd operative condition -- it differs from the case like. [ change / between the 1st and 2nd speech information trains / playback ]

[0038] For example, when a native speaker's "It's not much of a problem." is not able to be caught during playback of the 1st speech information train, "It is not much of a problem." which divides and speaks a word slowly can be heard by switching playback to the 2nd speech information train from the 1st speech information train under playback. And what is necessary is just to switch playback to the 3rd speech information train from the speech information train under playback further to know the semantics and the syntax of this Japanese. Of course, in this playback method, the thing which was explained by the playback method of the voice record medium concerning the 2nd embodiment of the above and to which it can apply so that it may return and can use combining a command or stop instruction cannot be overemphasized. Moreover, also in this playback method, change playback and repeat playback are possible, and specification of the target segment [ interrupt requests / these ] is performed like the case of the 1st embodiment mentioned above.

[0039] (The 4th embodiment of a voice record medium) the 4th embodiment of the voice record medium concerning this invention -- fundamental -- the 3rd operative condition of the above -- although it is the same as that of the case like, it mainly differs in that the alphabetic information train besides the 1st and 2nd speech information train is recorded. This alphabetic information train is equivalent to the alphabetic information corresponding to English (voice) which is equivalent to the alphabetic information of the contents corresponding to the 1st or 2nd speech information train, for example, a native speaker speaks.

[0040] This alphabetic information train is also classified into each segment of the 1st and 2nd speech information train, and a corresponding segment. moreover, this 4th operative condition -- also in the voice record medium applied like, the information about those head locations is included in record location identification information for every segment of each speech information train in the record location of this alphabetic information train, and it is recorded on the directory field of the voice record medium concerned. Moreover, boundary location identification information is recorded on the header unit of each segment. Therefore, the 1st and 2nd speech information train and an alphabetic information train will correspond per segment, respectively. in addition, this 4th operative condition -- the voice record medium applied like -- setting -- the 3rd above-mentioned operative condition -- when adding the 3rd speech information train which can be set like as recording information, 1 or two or more segments of the 1st and 2nd speech information train and an alphabetic information train will correspond also to

one segment group of the 3rd speech information train. Furthermore, also in this 4th embodiment, the head location of each segment is included in the above-mentioned record location identification information, and it is recorded on the directory field of the voice record medium concerned. Moreover, boundary location identification information is recorded on the header unit of each segment.

[0041] the playback method of the voice record medium (the 4th embodiment) equipped with the above structures -- fundamental -- the 2nd operative condition of the above -- although it is the same as that of the case like, it differs in that a display indication of the alphabetic information train is given during playback of the 1st or 2nd speech information train.

[0042] For example, when segment "It's not much of a problem." of the 1st speech information train is reproduced, a display indication of "It's not much of a problem." or the "It is not much of a problem." is given at a predetermined display. In addition, it is not necessary to synchronize with the speech information train and time amount target under playback completely about this display, and an alphabetic character may be displayed little by little behind time, or may be displayed little by little first. Moreover, also in this playback method, change playback and repeat playback are possible, and specification of the target segment [ interrupt requests / these ] is performed like the playback method of the voice record medium concerning the 1st embodiment mentioned above.

[0043] Next, the concrete structure of the voice record medium concerning this invention is explained to details below using drawing 2 - drawing 5.

[0044] Drawing 2 is drawing for explaining each speech information trains A, B, and C and the contents of record of those when applying the 3rd above-mentioned embodiment to English-conversation self-study as an example of the voice record medium concerning this invention. In this drawing, the speech information train A is an English information train (the 1st speech information train) about which a native speaker speaks, and consists of two or more segments 621 and 622. It is the information train (the 2nd speech information train) of English which divides a word and is spoken slowly, and, as for the speech information train B, this also consists of two or more segments which consist of the previous alphabetic word or previous phrase corresponding to segments 621 and 622 of the speech information train A. Moreover, the segment group which the speech information train C is an information train (the 3rd speech information train) which explains Japanese, and is contained in this speech information train C supports each segments 621 and 622 of each speech information trains A and B, respectively.

[0045] Moreover, drawing 3 is a table for explaining the time amount per one segment and the relation of capacity to the mode shown in drawing 2. In this table, the capacity of 6 K bytes is supported for 1 second. for example, in the segment 621 of the speech information train A The capacity for 0.2 seconds 1.2KB (kilobyte), [ the utterance time amount of "It's" ] The capacity for 0.1 seconds 0.6KB (kilobyte), [ the utterance time amount of "not" ] The capacity for 0.4 seconds 2.4KB (kilobyte), [ the utterance time amount of "much of a" ] And the utterance time amount of "problem" is 0.3 seconds, the capacity is 1.8KB (kilobyte), the utterance time amount of the segment 621 whole sets, and the capacity is set to 12KB (kilobyte) for 2.0 seconds. In addition, the numeric character of the parenthesis in a table is the value which showed the boundary location by playback time amount. That is, in the case of this speech information train A, since a boundary location is set up after "not" as shown in drawing 1 (a), it is shown that the boundary which divides the speech information train A into a portion a part for the first portion and the second half is set as the location 0.3 seconds after playback initiation of a segment 621.

Moreover, although other speech information trains B and C are shown similarly, in the embodiment of this drawing 3, boundary location identification information is not set to the speech information train C.

[0046] Furthermore, drawing 4 is a table for explaining the contents of record of the directory field in the mode shown in drawing 2 and drawing 3. In this table, a directory field consists of  $9 \times 3 = 27$  byte (B) per one segment. The speech information trains A, B, and C are equivalent to the speech information trains A, B, and C of drawing 2, respectively. Moreover, 1 byte of C shows an attribute and it means that  $C = 0$  is the speech information train A, and  $C = 64$  is the speech information train B. moreover,  $C = -$  - it means that 128 and 129 are the speech information trains C, and shows especially that it is the same candidate for description as a front segment at the time of "10000001" (belonging to the same segment

group set as the description object of the speech information train C is shown, for example, the case of the segments 801 and 802 of drawing 1 (e) corresponds), the time (8 bits (bit)), i.e., a bit expression, of  $C = 129$

[0047] M, S, and B (1 byte each) of positional information are a parameter which expresses with the industrial world the location on CD-ROM which is a standard. That is, in M, a part and S show a second and B shows a block, respectively. Moreover, 1 block is 2,048 bytes and constitutes 1 second from 75 blocks. Therefore, the maximum number is set to  $M = 59$ ,  $S = 59$ , and  $B = 74$ . 2 bytes of following SB shows a start cutting tool, and 3 bytes of the following LLL shows the length of each whole segment. In addition, in the parameter which shows a location, CD-ROM is because it was developed as an object for music from the first, and the reason using a part and a second expresses a record location as time amount from the start. Therefore, when CD-ROM is adopted as the voice record medium concerned, it will be only the information that at this rate and the second of the time amount at the time of playback are completely unrelated to each other, and only expresses the record location on a record medium.

[0048] Consequently, for example, "It's not much of a problem." of the segment 621 in the speech information train A The English speech information about which NEITIPUSUBIKA speaks by length of 3 blocks the 826th byte to 6,000 bytes for 0 minutes and 11 seconds is recorded. As for the corresponding segment in the speech information train B, English which divides a word and is spoken is slowly recorded by 0-minute and 11-second a length of 3 blocks the 2,026th byte to 17,400 bytes. As for the segment group of the speech information train C, Japanese description is recorded by length of 6 blocks the 1,282nd byte to 72,000 bytes for 0 minute and 11 seconds. In addition, there is no segment number of 621 and 622 grades on memory, and it supports the address. Moreover, the record location identification information which shows the relation of each segment is contained to this directory field.

[0049] The information for which an attribute C is equivalent to "It's not much of a problem." about which a segment speaks by 621, the speech information train, i.e., the native speaker, of 0, is still more specifically recorded on the field of a  $826 + 6,000 - 1 = 6,825$  byte up to [ from the contents of record of the directory field shown in drawing 4 / from the 826th byte in 3 blocks (0 minute and 11 seconds) of the voice record medium concerned ]. Moreover, the information from which a segment is equivalent to "It is not much of a problem." about which an attribute C divides, the speech information train, i.e., the word, of 64, and speaks slowly by 621 is recorded on the field of a  $2,026 + 17,400 - 1 = 19,425$  byte up to [ from the 2,026th byte in 3 blocks (0 minute and 11 seconds) of the voice record medium concerned ]. Furthermore, the information from which an attribute C is [ a segment ] equivalent to the speech information train of 128, i.e., Japanese description, by 621 is recorded on the field of a  $1,282 + 72,000 - 1 = 73,281$  byte up to [ from the 1,282nd byte in 6 blocks (0 minute and 11 seconds) of the voice record medium concerned ].

[0050] Thus, if the directory field shown in drawing 4 is prepared, each speech information train shown in drawing 2 by playback time amount and capacity as shown in drawing 3 is recordable.

[0051] Next, the boundary location identification information which shows the boundary of a portion in a part for the first portion and the second half of each segments 621 and 622 is recorded on the header unit of the variable-length segment shown in drawing 5 (a). A 1-byte field to show the existence of alphabetic information or image information etc. from a head, as this header unit was shown in drawing 5 (b) (1B), the field prepared for speech information train A -- it is -- an information train type (the speech information train A --) The 1-byte data in which the information for distinguishing B etc. is shown, the 3-byte data in which the data length is shown, And the 5-byte field which consisted of 1-byte data in which a boundary location is shown (5B), The 1-byte data in which it is the field prepared for speech information train B, and an information train type is shown, The 5-byte field which consisted of 3-byte data in which the data length is shown, and 1-byte data in which a boundary location is shown (5B), The 4-byte field which consisted of 3-byte data in which the 1-byte data in which it is the field prepared for speech information train C, and an information train type is shown, and its data length are shown (4B), The 4-byte field which consisted of 3-byte data in which the 1-byte data in which it is the field prepared for alphabetic information-train D, and an information train type is shown, and its data length are shown (4B), The 6-byte field which consisted of 3-byte data in which the 3-byte data in which

it is the field similarly prepared for alphabetic information train D, and the address is shown, and its data length are shown (6B), It is 32 bytes of field which consists of a 4-byte field (4B) prepared for other information trains (Type E), and a spare 3-byte field (3B).

[0052] In addition, other information trains (Type E) are not used in the embodiment explained in this specification. Moreover, in this embodiment, boundary location identification information is set only to the speech information trains A and B. The quotient which divided the byte count from a segment head by 256 in the information after compression is set to the 1-byte field which shows a boundary location. For example, if the voice sampling rate of A truck (virtual truck equivalent to the speech information train A) is made into 128 kilobits per second (16 K bytes/(second)), the boundary location of the range from a head to 4 seconds can be expressed in this 1-byte field, and that resolution will become 0.016 seconds. Moreover, when voice sampling rates are 32 kilobits per second (4 K bytes/(second)), the boundary location of the range from a head to 16 seconds can be expressed in this 1-byte field, and that resolution becomes 0.064 seconds, and practically sufficient engine performance is obtained.

[0053] Next, the fundamental playback method of the voice record medium concerning this invention is explained using drawing 6 - drawing 10.

[0054] First, drawing 6 is the perspective diagram showing the whole regenerative-apparatus configuration for [ of the voice record medium concerning this invention ] carrying out the playback method implementation. As shown also in this drawing, remote control of the voice record medium concerned is carried out with the hand set 8 which is refreshable CD-ROM and by which code connection of this main part 2 of a regenerative apparatus was made with the portable CD player (main part 2 of a regenerative apparatus). The displays 21, such as a liquid crystal display (LCD) which displays the segment number under playback at least, and the various carbon button groups 240 for control are formed in this hand set 8. Moreover, an operator can hear the speech information reproduced by the main part 2 of a regenerative apparatus through an earphone 13.

[0055] Moreover, drawing 7 is the block diagram showing the configuration of the regenerative apparatus shown in drawing 6. As shown in this drawing, CD-ROM which is the voice record medium 1 concerned is set to the playback device 20. The playback device 20 is connected to CPU5 through disk interfacing (I/F) 3 and a bus 4. Moreover, ROM6 of 32 K bytes (KB) for storing a program and 256 K bytes of RAM7 for storing a directory and a speech information train temporarily are connected to the bus 4. Furthermore, it connects with the bus 4 at D/A KOMPATA 12 connected to the external terminal 11 and the hand set 8 through the hand-set interface (I/F) 9 which delivers and receives information between the hand sets 8 for manual operation, and the amplifier (AMP) 10 for voice outputs. In addition, the earphone 13 is connected to the hand set 8 as mentioned above.

[0056] Drawing 8 (a) and (b) are drawings for explaining the memory quota condition of ROM6 and RAM7, respectively. A program is stored in 32 K bytes of ROM6 as shown in drawing 8 (a). On the other hand, as shown in drawing 8 (b), a  $= (50+50)$  100 K byte buffer (equivalent to 50 blocks), a  $= (75+75)$  150 K byte directory, and the system area for 6 K bytes are assigned to RAM7. Therefore, the speech information train for 50 blocks is always held at RAM7, and the directory for  $27 \times 150$  K byte / 5,555 segment (considerable [ in about 30 minutes ] only in the portion of the speech information train A) is held.

[0057] In addition, although CD-ROM is used as the voice record medium concerned by the above-mentioned example, the capacity of the typical thing is 552 megabytes (MB). The part, the second, and the unit of a block are used for expressing the address with CD-ROM. Moreover, since 1 block is [ 1 second and 60 seconds of 2,048 bytes and 75 blocks ] 1 minute, the greatest value of the address of this CD-ROM is 74 blocks (59 minutes and 59 seconds). On the contrary, the capacity of this CD-ROM is  $2,048 \times 75 \times 60 \times 60 = 552.96$  megabyte. Among these, as a format of CD-ROM, since the amount of 2 seconds after the beginning cannot employ a user, it considers as maximum capacity correctly, and it is set to 552.6528MB. Furthermore, if a directory field is assigned from the beginning to equivalent to 20 seconds, then time, the directory capacity of 3 megabytes is securable for CD-ROM.

[0058] Next, the example of count about capacity is shown.

[0059] Make an audio sampling rate into the 16K sample / second by the ADPCM method, and let it be



a triplet per one sample. Thus, if it assumes, the amount of voice samples needs to become a second in 48 kilobits per second (Kbit/s) (KB/s) /, i.e., 6 K bytes, and also needs to double a transfer rate with this. In addition, if it is a 16K sample / second, there is an f property to 8kHz (kHz), and it can fully record to a consonant. Thus, when it assumes, 6 K-byte x3,600 second =21.6 megabyte capacity is needed for the sound recording of 1 hour. Generally, an error collection is put into one CD-ROM, and 552 megabytes can be recorded on it. If a directory field is removed, 549 megabytes can be used as each data area which records the account of \*\*\*\*\*. Therefore, it becomes possible to record the speech information for 549/ 21.6= 25 hours and 24 minutes. Then, if it makes into 1 hour the tale pronounced at a natural speed by the native speaker in using as an object for English conversations, the portion which divided word each and has pronounced it slowly can be put at about 4 hours by the 4 times. And even if there are 15 hours of portions of description in total, they are 20 hours in all.

[0060] Next, considering into how many segments (partition) the conversation portion of 1 hour can be decomposed, they are 14,400 pieces as an average of four breaks in 1 second in 1 hour. Since it is 36-byte necessity at DIKUTORI of one segment, it is about 518 K bytes on the whole, and this can be enough contained in the receipt location of the above-mentioned 3 megabytes of directory, and can store all the directories of the tale of 1 hour.

[0061] Next, drawing 9 is the plan showing concretely the configuration of the hand set 8 and earphone 13 which were shown in drawing 6 and drawing 7. The display 21 (LCD) which displays a segment number on the front face of a hand set 8, In order to switch the display 22 and playback mode (playback conditions in case repeat playback is carried out are shown) which show the contents of directions at present using LED A mode switch 81, each speech information train A The command switches 23A, 23B, and 23C for carrying out the playback command of B and C, REV carbon button 24B for carrying out ADV carbon button 24A for carrying out functional directions, and a reproductive return command, The volume switch 28 for adjusting STOP carbon button 25B for ordering it a halt of REP carbon button 25A for directing repeat playback and playback and the sound volume outputted from an earphone 13 is formed. Moreover, an earphone 13 is connected to a hand set 8 through a code 26, and the hand set 8 is connected to the main part 2 of a regenerative apparatus through the code 27.

[0062] Next, an example of the fundamental playback sequence in the playback method of the voice record medium concerning this invention is explained. Drawing 10 is drawing for explaining this fundamental playback SHIKEN. When continuing and hearing the portion of the speech information train A, i.e., a native speaker's English, the speech information of the segment by which sequential record was carried out as shown in this drawing 10 (a) is reproduced. in this case, carbon button 23A shown in drawing 9 is pushed, and it is heard as it is -- being sufficient.

[0063] On the other hand, in the playback sequence shown in drawing 10 (b), playback of the speech information train A is first performed sequentially from a segment 621. Here, when a segment 623 is not able to be caught well, an operator once interrupts voice playback by pushing STOP carbon button 25B immediately. At this time, the segment number of the segment display 21 is 624. Then, REV carbon button 24B is pushed only once. In addition, if it is continuing pushing REV carbon button 24B, whenever it will return continuously and will push once, every one segment number displayed on the display 21 returns. The segment number as which the segment was displayed on the display 21 in the condition of 624 when REV carbon button 24B was pushed once is set to 623.

[0064] Then, when switch 23B is pushed (change playback directions), the voice which the segment 623 of the speech information train B corresponding to the segment 623 of the speech information train A was reproduced, divided each word, and has been pronounced slowly is \*\*\*\*\*. If left as it is, playback actuation of a segment 624 will advance continuously with the speech information train B. Furthermore, after STOP carbon button 25B is pushed in the place where the segment 624 of this speech information train B was heard, if REV carbon button 24B is pushed again, the segment number of a display 21 will be set to 623. If switch 23C is pushed in this condition, description of Japanese of the speech information train C will be reproduced. Since the portion of description may explain the phrase some whose words settled, the thing (an attribute C is the thing of 129) of 1 shows it, and the least significant bit of the cutting tool whom one description also shows a certain attribute C with some segment

numbers means that it is the same candidate for description as the segment number in front of the same speech information train. It can skip in here.

[0065] Next, other embodiments of the voice record medium concerning this invention are explained using drawing 11 - drawing 13.

[0066] Drawing 11 is drawing showing other examples of a configuration of the directory field in the voice record medium concerning this invention. The directory shown in this drawing consists of 15 bytes per one segment. That is, it is a total of 12 bytes of the attribute C of 1 byte, 2 bytes of segment number SS, 1 byte of suffix number N, 1 byte of part M, 1 byte of second S, 1 byte of block B, and 3 bytes of segment length LLL.

[0067] Moreover, in 8 bits of an attribute C, when the 1st bit (most significant bit) is 1, the start of a segment is meant, and it means that it is in other conditions at the time of 0. Moreover, when the value expressed by the 2nd triplet in this attribute C is 0, it means that it is the speech information train A, and it means that it is the speech information train B at the time of 1, and means that it is the speech information train C at the time of 2. Furthermore, when the 4th bit is 0, it means that there is no suffix, and it means that there is a suffix at the time of 1. When the 5th bit is 1, it means being related with the same segment group as a front segment, and it means not being related with the same segment group at the time of 0.

[0068] Drawing 12 is drawing for explaining the division condition of the segment of each speech information trains A, B, and C corresponding to the directory shown in drawing 11. And the point that this differs from the example shown in drawing 2 is a point classified by the segment which attached the suffix.

[0069] Moreover, drawing 13 is drawing showing the relation between the value of the 2nd triplet with an attribute [ C ] of 8 bits, and the segment equivalent to each divided partition, as explained above.

[0070] furthermore, other operative conditions of the voice record medium applied to this invention using drawing 14 and drawing 15 -- it needs -- it explains. In addition, the point that other embodiments differ from the embodiment of drawing 11 explained previously - drawing 13 is a point that this embodiment is made to perform not only speech information but record of alphabetic information and playback.

[0071] Drawing 14 is drawing for explaining the structure of the directory. In this drawing, a different point from the directory of drawing 11 explained previously is a point that the information about the alphabetic information train D is added to the 2nd and 3 bit of an attribute C. When this bit is 3, the alphabetic information train D means what is recorded on the voice record medium concerned in predetermined code.

[0072] Drawing 16 is the plan showing the configuration of the hand set used for the regenerative apparatus of the voice record medium concerning this embodiment. In addition, the difference between the hand set of this drawing 15 and the hand set of drawing 9 is a point which is the mode which can be displayed also about the alphabetic information train D. It is the point that the character representation section 210 concretely constituted from LCD by the hand set 8 with manual operation button 23D for directing character representation is formed. Moreover, in this embodiment, the character representation carbon button for making the display of an alphabetic character turn on and turn off is also prepared further. In such a hand set 8, a setup of being \*\*\*\*\* is possible in the alphabetic information train D in the character representation section 210.

[0073] Next, the special feature of an operation of other embodiments shown in drawing 11 - drawing 13 is explained.

[0074] The speed of the pronunciation of the information train A, i.e., the speed in which that alphabetic information train D appears according to the length of LLL, is controlled by this embodiment during playback of the speech information train A. That is, an alphabetic character appears in a display 210 and it is made to finish outputting completely to from the beginning of the pronunciation of this segment before the end (it is made to synchronize with pronunciation completely and an alphabetic character is displayed).

[0075] Next, during playback of the speech information train B, since it is pronounced over several



times as many time amount as this compared with playback of the portion of the speech information train A, it is made to synchronize with the length of the pronunciation of the speech information train B at this time, and the alphabetic information train D is displayed. What is necessary is just to set this point as arbitration, since it is considered that it is also convenient for those whom how to issue the display of an alphabetic character a little early (a little voice output is delayed and taken out) employs in fact.

[0076] Next, the case where it considers as the magnetic tape to which an information record medium is not made as for random access as an embodiment shown in drawing 14 and drawing 15 is explained.

[0077] this operative condition -- although it also sets like and record of fundamental information and a reproductive mode are the same as that of an above-mentioned embodiment -- the MAG -- in order to consider as a practical regenerative apparatus since jump playback of data is not easy if deep, it is desirable to form the buffer for memorizing information temporarily. Specifically, it can stand in a line in order considering each segment of the speech information trains A, B, and C which corresponds each subunit equivalent to the speech information trains A, B, and C (three trucks) of the above-mentioned virtual truck like A1, B1, C1, A2, B-2, C2, A3, B3 and C3, and -- as one lump. And when taking out and reproducing only the segment of the speech information train A like A1, A2, and A3 when reproducing the speech information train A, and reproducing the speech information train B, only the segment of the speech information train B is taken out and reproduced like B1, B-2, B3, and --. The same is said of the speech information train C. At this time, in order to reproduce a segment A2 without a break to the degree of a segment A1, it must fly and the segments B1 and C1 which exist among these segments must be exceeded. To the usual tape transit, since this is not easy, it takes information in the buffer a little early from the magnetic tape. And it is made to perform jumping and reproducing in a buffer.

[0078] It is as follows when this magnetic tape is set to the so-called DAT (digital audio tape). First, a native speaker's pronunciation recorded on A truck is divided in a certain unit as a segment of the speech information train A. For example, it will divide into fixed time amount in every second here. and as a truck B (speech information train B), it considers as the thing of the contents about which it is speaking at one third of speed an average of [ for example, ] -- slowly. A field is secured so that A truck's being equivalent to 1 second, then time can be taken by 4 seconds on B trucks. And a field is secured so that it can take by 9.4 seconds as a C truck for description. In addition, the above-mentioned number of seconds is an example, and is limited to this. Then, there is nothing. That is, let 14.4 seconds be one unit. That is, three segments to which A, B, and C truck (equivalent to each speech information trains A, B, and C) each correspond are put together for 14.4 seconds. Are only cutting the original thing by a unit of 1 second, and the amount of [ \*\* which needs cautions here, and for A trucks ] 1 second is completely in agreement with 1 second of a basis. That what is necessary is just less than 4 seconds, it has finished in 3 seconds and 4 seconds of B truck may be as \*\*\*\* for 1 second. In addition, about the method of flying, when reproducing this surplus place (1 second), it mentions later.

[0079] In the case of the rotary head mold DAT, it is as follows. That is, the method which records or reproduces two trucks by one rotation of a rotary head in this rotary head mold DAT is common. And 2,880 bytes of speech information can be recorded now on one truck. Therefore, it will become 86,400 bytes if 30 trucks are cut off the 1st division. The sampling rate of the voice explained also in the above-mentioned embodiment into this = when voice is recorded by 48 (Kbit/s), it is 86,400(cutting tool) / 6,000(cutting tool) = 14.4 (s).

\*\*\*\* record can be carried out.

[0080] Next, a buffer is considered.

[0081] Since it will be set to  $10,480,576 / 86,400 = 12,135$  if 1MB of RAM is used, 12 buffers can be formed. Twelve buffers can be stood in a line in the shape of a ring in functional semantics, and it always prefetches two information by segment at a time, and transmits to a buffer from a tape. Playback is performed using the information transmitted to this buffer. When it stops in the place of arbitration, the voice before the time of ten buffers stopping (what has already been taken out from the loudspeaker as a playback sound) remains. That is, repeat playback can be carried out from a buffer, without there being voice on A truck and the voice for 10 seconds rewinding a tape.

[0082] By doing in this way, it can move to every virtual truck of A, B, and C easily in the segment of

arbitration as well as the example using CD-ROM mentioned above about the speech information for [ it remains in ten buffers ] 10 seconds.

[0083] However, since random access is impossible like CD-ROM, the number of time amount of the segment to which three trucks, A, B, and C, correspond must always carry out within 14.4 seconds. At this time, the portion of the virtual truck A may be mechanically divided so that it may always become fixed time amount width of face (every [ for example, ] second). For example, since the portion of B truck is also at the time of 2 seconds and is also at the time of 5 seconds, it is made to become adjustable. The time amount which lengthened the time amount of A and B truck from 14.4 seconds is the time amount assigned to C truck. In addition, the portion which corresponded to A truck in a break completely did not need to be recorded, it let two or more nearby breaks pass, and both B and C truck have just performed a certain settlement.

[0084] Since there is a sub-code field in DAT, the information on a byte count which shows the boundary line of A in each break (every 14.4 second / 30 truck) and each break, B, and C truck is recordable on this sub-code field.

[0085] Next, the amount of speech information recordable on DAT is calculated concretely.

[0086] Generally, 240,000 trucks are recordable on DAT for 1 hours. The six next trucks are used for the truck of a rest, and when having stopped the tape, it is made for a rotary head not to give a crack to the portion on which speech information is recorded as the connoisseur many times, although 30 trucks are made into 1 break. 36 trucks become one unit in order to do in this way. That is, on the whole, 6,666 units can be recorded and it has been 126 hours, 39 minutes,  $6,666 \times 14.4\text{-second} = 50\text{ seconds}$ . Among these, since it assigns as a part for A truck in 1 hour, 50 seconds per part can be spent on B and C truck for 39 seconds for 25 hours. Even if B truck pronounced slowly uses by 4 hour, it can assign C truck for explanation for 21 hours or more, and sufficient length can be secured.

[0087] In addition, how to take the segment of the above-mentioned A, B, and C truck is the same with the thing adapting the above-mentioned CD-ROM, and A truck may make a second degree a standard and it not only divides it by fixed time amount width of face, such as 1 etc. second, but it may cut it in a place with the end eye on a certain pronunciation. That is, you may change with the editorial policies when recording. Therefore, it is also possible to also divide that whose random access is possible like CD-ROM by fixed time amount width of face, and it to classify a segment.

[0088] Next, interrupt processing about the playback method especially a repeat playback demand, and change playback of the voice record medium concerning this invention is explained using drawing 16 - drawing 18.

[0089] Drawing 16 is a timing diagram for explaining decision actuation of the segment for interruption in the playback method of the voice record medium concerning this invention.

[0090] Usually, by the playback method of the voice record medium concerning this invention, the segment (--, N-1, N, N+1, --) of the speech information train directed beforehand is reproduced in order. It is in the middle of this playback, and when the interrupt request (R1, R2, R3, and -- show among drawing) about repeat playback, change playback, etc. occurs from operators (for example, English student etc.), the segment set as the object of an interrupt request is specified first. That is, as shown in drawing 16, when Segment N is reproduced, a front [ location / SN / where the starting interrupt request is shown by the boundary location identification information recorded on the header unit of each segment explained previously / boundary ], or the back is judged. In addition, the boundary location in each segment is shown by --, SN-1, SN, SN+1, and -- all over drawing, respectively. furthermore, the segment N which was under playback as an object segment of the interrupt request concerned based on the decision result about the generating timing of this interrupt request -- or the segment N-1 before that is specified. By specifying the object segment of the generated interrupt request, playback actuation of speech information is resumed according to the contents of directions of this interrupt request from the head of the segment of other speech information trains corresponding to the head of this specified segment, or the specified this segment.

[0091] When the desired interrupt request R1 occurs during playback of the segment N of a certain speech information train, it means generating the generating timing of this interrupt request R1 in a part

for the first portion of Segment N concretely. In this case, the segment N-1 in front of [ of the segment N which judged to be an operator's directions delay and was under playback ] one is specified as an object segment of an interrupt request. On the other hand, when an interrupt request R2 occurs in the second half portion of the segment N which was under playback, the segment N itself which was [ this ] under playback is specified as an object segment of an interrupt request. In addition, in the interrupt request R3 generated during playback of the segment N+1 shown, other segments, for example, drawing 16, the generating timing of an interrupt request is similarly judged on the basis of boundary location SN+1, and the object segment of an interrupt request is specified.

[0092] Drawing 17 is a flow chart for explaining repeat playback interrupt processing (the 1st embodiment concerning the playback method) in the playback method of the voice record medium concerning this invention.

[0093] In the voice record medium prepared by this playback method It is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means at least. The speech information train which was divided at the turning point of pronunciation, respectively and which was classified into the variable-length segment for every speech information, The boundary location identification information which shows a boundary with a portion in a part for the first portion and the second half of this each partition in this speech information train divided by the pronunciation condition for every partition should just be recorded on the desired field, respectively.

[0094] First, if the interrupt request about repeat playback occurs, while the information on the segment which was once under playback (segment number) will be held by the predetermined maintenance means (step ST 1), the information about the generating location of the interrupt request concerned is also once held [ be / it / under / playback / of the above voice record media / setting ] (step ST 2). And the boundary location and interruption generating location which were under playback and which were recorded on the header unit of a segment are compared along with a time-axis (SUTEPU ST 3). When an interrupt request occurs after a boundary location at this time, the segment itself which was under playback is specified as an object segment of the interrupt request concerned, reading appearance of the head positional information of the specified segment concerned is carried out from a directory, and voice playback actuation is resumed (step ST 4). On the other hand, when an interrupt request occurs before a boundary location, from the segment which was under playback, the segment in front of one is specified as an object segment of the interrupt request concerned, reading appearance of the head positional information of the specified segment concerned is carried out from a directory, and voice playback actuation is resumed (step ST 5).

[0095] In addition, in this embodiment, repeat playback of the speech information of the specified segment is carried out until the directed repeat playback actuation has the following new directions.

[0096] Next, change playback interrupt processing (the 2nd embodiment concerning the playback method) in the playback method of the voice record medium concerning this invention is explained using the flow chart of drawing 18.

[0097] The voice record medium prepared in this embodiment is a speech information train corresponding to 1 or two or more sentences which consisted of two or more word trains by which a playback output should be carried out with a predetermined voice playback means. The speech information train A which was divided at the turning point of pronunciation, respectively and which was classified into the variable-length segment for every speech information The speech information train B which is a speech information train which consisted of pronunciation which was equivalent to this speech information train A, and \*\*\*\*(ed), and was divided into the variable-length segment for every speech information equivalent to the speech information of each segment of the speech information train A The record location identification information which shows each segment which can be switched in the record location of each of this segment in the record medium concerned that the speech information trains A and B should be switched and it should reproduce, And the boundary location identification information which shows a boundary with a portion a part for the first portion (judged as the settlement of pronunciation) divided according to the pronunciation condition for every segment and the second

half in the speech information trains A and B should just be recorded at least.

[0098] First, if the interrupt request about change playback occurs, while the information on the segment which was once under playback (segment number) will be held by the predetermined maintenance means (step ST 6), the information on the newly directed speech information train is held [ be / it / under / playback / of the above voice record media / setting ] (step ST 7). Moreover, the information about the generating location of the interrupt request concerned is also once held (step ST 8). And the boundary location and interruption generating location which were under playback and which were recorded on the header unit of a segment are compared along with a time-axis (SUTEPU ST 9). When an interrupt request occurs after a boundary location at this time, the segment itself which was under playback is specified as an object segment of the interrupt request concerned, it is the segment of the speech information train directed by the interrupt request, and reading appearance of the head positional information of the this specified segment is carried out from a directory, and voice playback actuation is resumed (step ST 10). On the other hand, when an interrupt request occurs before a boundary location, from the segment which was under playback, the segment in front of one is specified as an object segment of the interrupt request concerned, it is the segment of the speech information train directed by the interrupt request, and reading appearance of the head positional information of the segment corresponding to the specified segment concerned is carried out from a directory, and voice playback actuation is resumed (step ST 11).

[0099] In addition, this invention is not limited to each above-mentioned embodiment, and can deform variously.

[0100] For example, linkage with a video disc or a video tape can also apply this invention. That is, the sound track section of a movie is put into A truck of an ordinary speed (a native speaker's utterance speed), the pronunciation carried out slowly is put into B truck, and Japanese description is put into C truck. And reproducing A truck, it stops in the place which was not able to be understood, a few is returned (at this time, a screen is equal to practical use also with immobilization), a \*\* B truck is heard, and if it still does not understand, C truck is heard. In addition, what is necessary is to make it synchronize with a screen, only when reproducing A truck.

[0101] Moreover, invention of smell levers, such as a personal computer, can also be used. That is, more advanced application is possible if it combines also with the programmable thing of linkage with CAI (Computer Aided Instruction), or a CDI (Compact Disk Interactive) equal altitude. Moreover, it is also possible to use the storage of a personal computer instead of the above-mentioned CD or DAT, and to carry out this invention.

[0102] Furthermore, it can apply not only to English conversation etc. but to the thing for \*\*\*\* and laws, and an information train may be three sorts or four sorts of not only things but more than it.

---

[Translation done.]

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

---

## DESCRIPTION OF DRAWINGS

---

### [Brief Description of the Drawings]

[Drawing 1] It is drawing for explaining notionally the speech information \*\*\*\*\* various information which should be recorded on the voice record medium concerning this invention.

[Drawing 2] It is drawing for explaining each speech information train and its contents of record of the voice record medium (the 3rd mode) applied to the English-conversation self-study concerning this invention.

[Drawing 3] It is a table for explaining the relation between the time amount per each segment, and capacity about each speech information train shown in drawing 2.

[Drawing 4] It is a table for explaining the contents of record of the directory field in the voice record medium shown in drawing 2 and drawing 3 (record location identification information being included).

[Drawing 5] It is drawing showing the configuration (boundary location identification information is included) of the variable-length segment which should be recorded on the voice record medium concerning this invention.

[Drawing 6] It is the perspective diagram showing the whole regenerative-apparatus configuration which realizes the playback method of the voice record medium concerning this invention.

[Drawing 7] It is the block diagram showing the configuration of the regenerative apparatus shown in drawing 6.

[Drawing 8] It is drawing for explaining the memory quota condition of ROM and RAM which were shown in drawing 7.

[Drawing 9] It is the plan showing the hand set and the whole earphone configuration which were shown in drawing 6.

[Drawing 10] It is drawing for explaining the fundamental playback sequence in the playback method of the voice record medium concerning this invention.

[Drawing 11] It is drawing showing other examples of a configuration of the directory field in the voice record medium concerning this invention.

[Drawing 12] It is drawing for explaining the division condition of the segment of each speech information train corresponding to the contents of record of the directory field shown in drawing 11.

[Drawing 13] In the division condition shown in drawing 12, it is drawing for explaining the relation between the 2nd and the value of the 3rd bit in the attribute C of 8 bits, and each divided segment among the contents of record of a directory field.

[Drawing 14] It is drawing showing other examples of a configuration of the directory field in the voice record medium concerning this invention.

[Drawing 15] It is the plan showing the outline configuration of a hand set of the regenerative apparatus which reproduces the voice record medium which has the directory shown in drawing 12.

[Drawing 16] It is drawing (timing diagram) for explaining decision actuation of the segment for interruption in the playback method of the voice record medium concerning this invention.

[Drawing 17] It is a flow chart for explaining the interruption processing about repeat playback in the playback method of the voice record medium concerning this invention.

[Drawing 18] It is a flow chart for explaining the interruption processing about change playback in the playback method of the voice record medium concerning this invention.

[Drawing 19] It is drawing (timing diagram) for explaining repeat interruption actuation in the playback method of the conventional voice record medium.

[Description of Notations]

1 [ -- An earphone, 21,210 / -- A display, 23A, 23B, 23C, 23D / -- A playback information train directions switch, 25A / -- Repeat directions switch. ] -- A record medium, 2 -- A regenerative apparatus, 8 -- A hand set, 13

---

[Translation done.]

## \* NOTICES \*

Japan Patent Office is not responsible for any damages caused by the use of this translation.

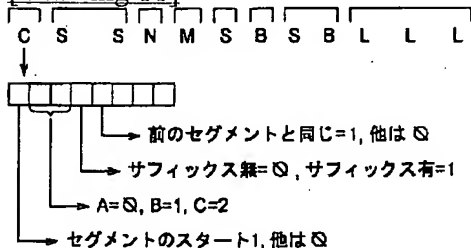
1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. \*\*\*\* shows the word which can not be translated.
3. In the drawings, any words are not translated.

## DRAWINGS

[Drawing 3]

セグメント			621	622	
時間 (秒)	A		2.0 (0.3)	1.3 (0.4)	
	B		3.9 (1.1)	2.6 (1.4)	
	C		12	25	
容量 (KB)	A		12	7.8	
	B		23.4	15.6	
	C		72	150	

[Drawing 11]



[Drawing 1]

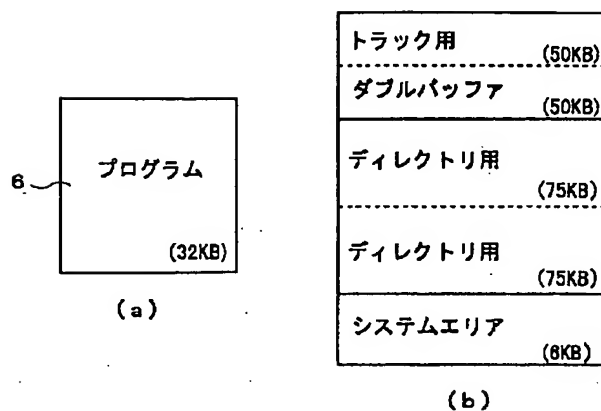
- セグメント 621
- (a) It's not <sup>↑</sup> much of a problem.  
S1  
セグメント 622
- (b) I'd second <sup>↑</sup> that.  
S2  
セグメント 701
- (c) Yes! <sup>↑</sup>  
S3  
セグメント 799
- (d) I'm gonna go fishing <sup>↑</sup> with my brother.  
S4  
セグメント 801
- (e) I'm gonna go fishing <sup>↑</sup> with my brother, ---  
S5
- セグメント 802
- because he'd bought a new rod <sup>↑</sup> yesterday.  
S6

[Drawing 2]

記録内容	
A	<div> <div>"It's not much of a problem."</div> <div>"I'd second that."</div> </div> <div> <div>621</div> <div>622</div> </div>
B	<div> <div>"It is not much of a problem. I would second that."</div> </div> <div> <div>621</div> <div>622</div> </div>
C	(621) 「It」とは前の〇〇を指しており、「not much a problem」は慣用句で『問題がない』とか『心配ない』というような意味です。
	(622) 次に、「I'd」とは「I would」をつめた言い方であり、同じように「I could」も「I'd」とつめて言えます。 「second」とは『支持する』、『賛成する』という意味があり、「I'd second that」で『同感です』という意味になります。

[Drawing 8]

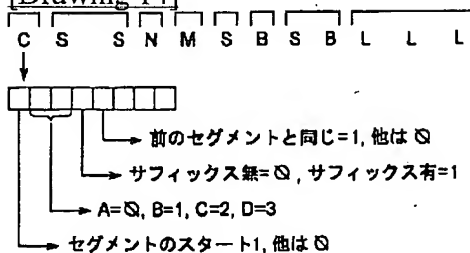




[Drawing 13]

2, 3ビット目	セグメント
0	6 2 1
1	6 2 1-1
1	6 2 1-2
2	6 2 1-1
2	6 2 1-2
2	6 2 1-3
2	6 2 1-4
0	6 2 2
1	6 2 2
2	6 2 2-1
2	6 2 2-2

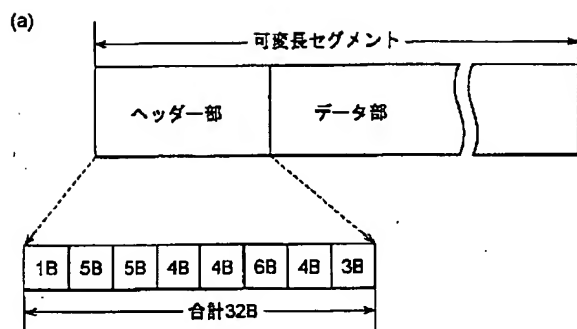
[Drawing 14]



[Drawing 4]

情報列	セグメント SS	カテゴリー C	位置情報 M S B			スタートバイト SB	セグメント量 (Byte) LLL
A							
A	621	0	0	11	3	826	6,000
B		64	0	11	3	2,026	17,400
C		128	0	11	6	1,282	72,000
A	622	0	0	11	49	2,018	4,800
B		64	0	11	50	570	12,600
C		128	0	11	52	1,274	160,000

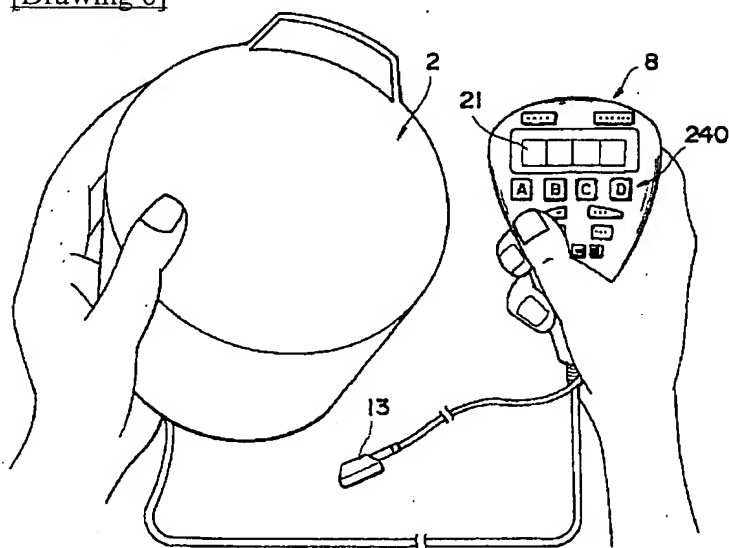
[Drawing 5]



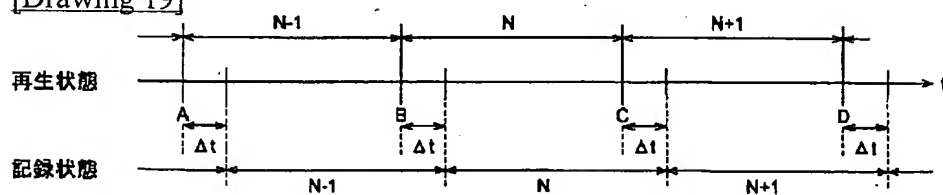
(b)

ヘッダー (合計32B)	情報列	内 容
1B		文字情報や画像情報の有無等 (1B)
5B	A	情報列タイプ (1B), データ長 (3B), 境界位置 (1B)
5B	B	情報列タイプ (1B), データ長 (3B), 境界位置 (1B)
4B	C	情報列タイプ (1B), データ長 (3B)
4B	(D)	情報列タイプ (1B), データ長 (3B)
6B	D	アドレス (3B), データ長 (3B)
4B	E	
3B	予備	

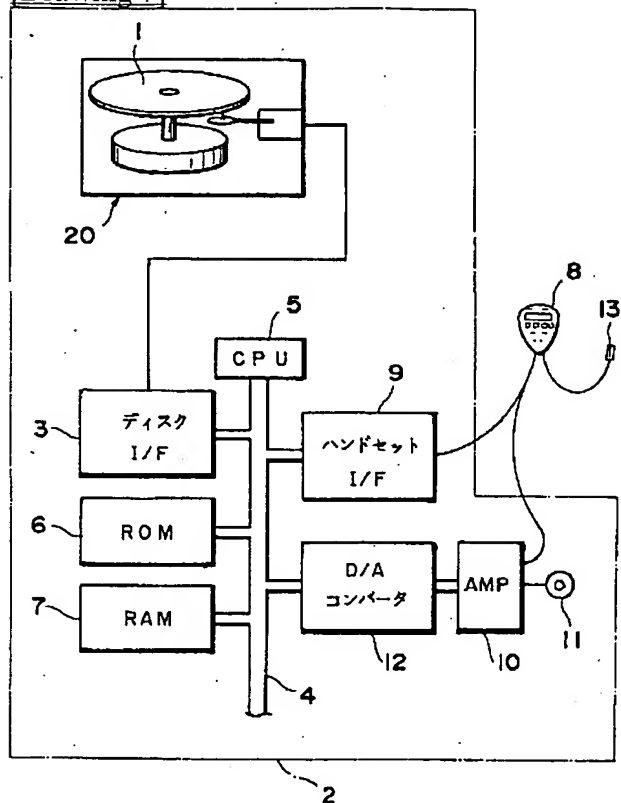
[Drawing 6]



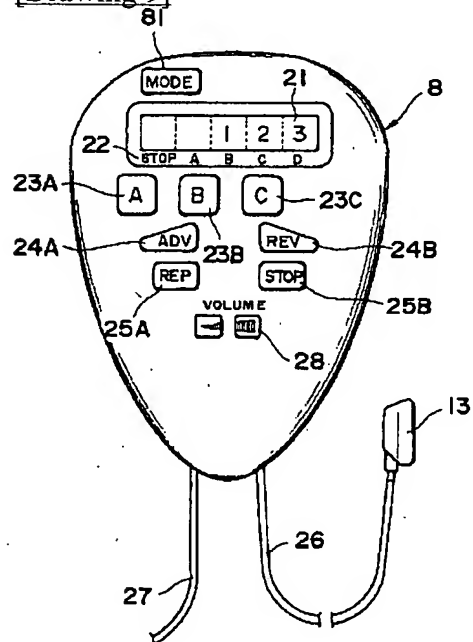
[Drawing 19]



[Drawing 7]

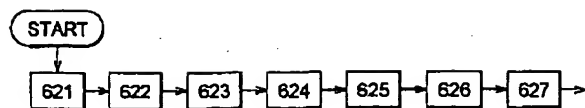


[Drawing 9]

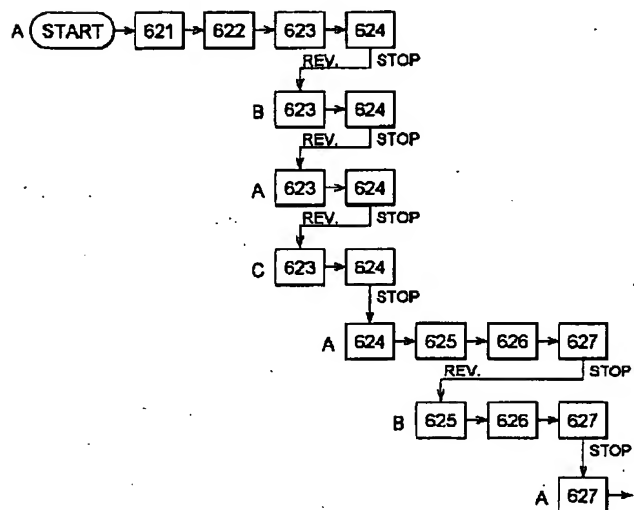


[Drawing 10]

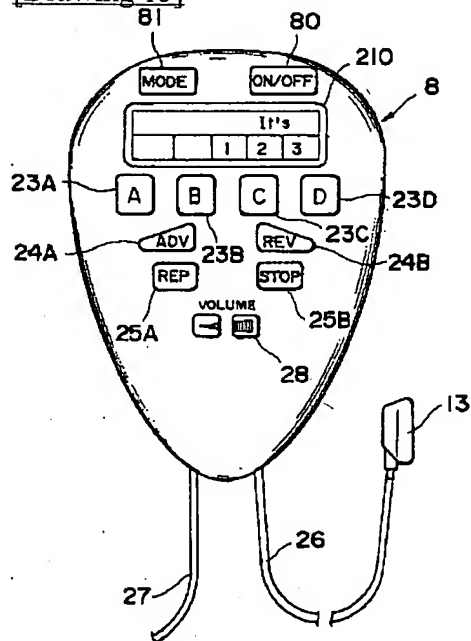
(a)



(b)



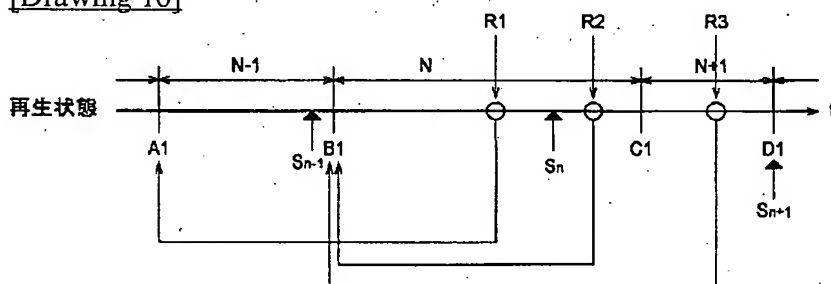
[Drawing 15]



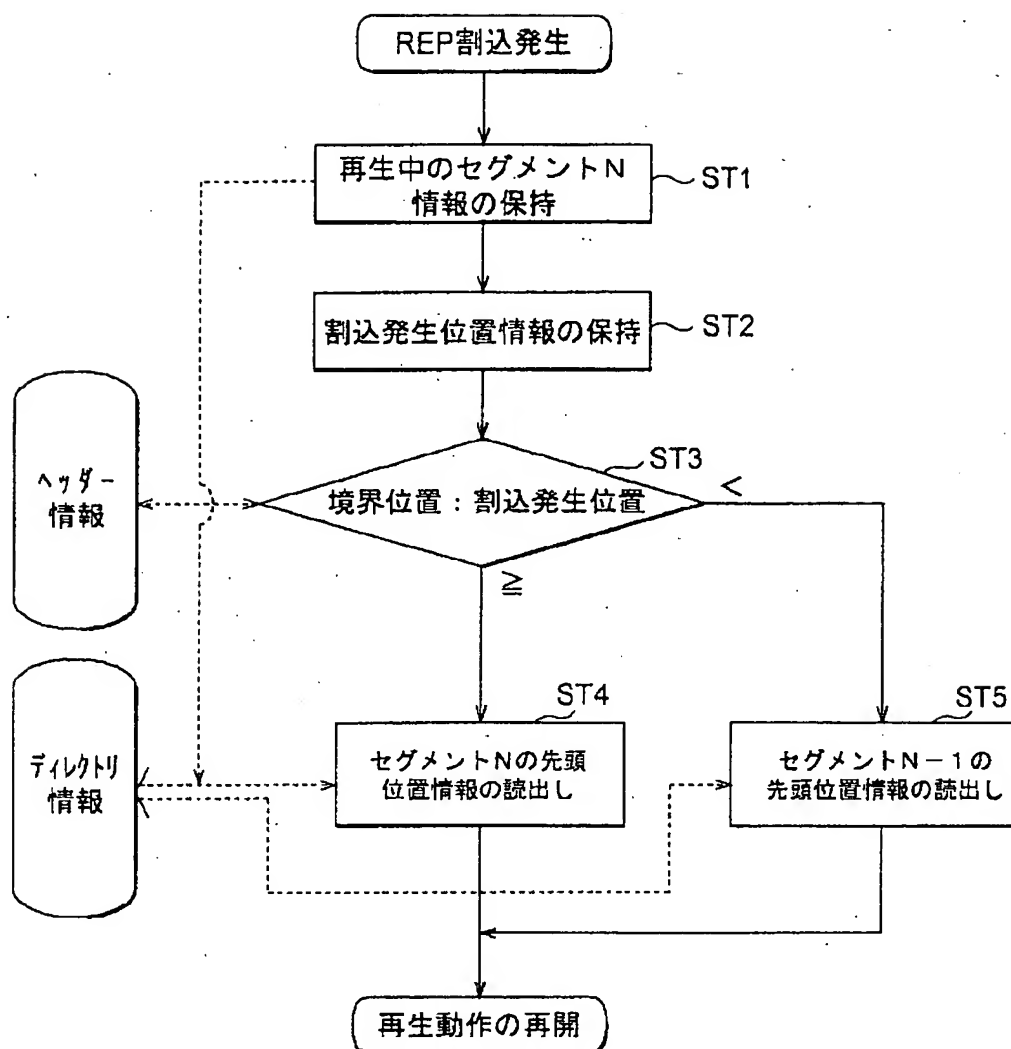
[Drawing 12]

情報列	セグメント分割
A	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; padding: 2px; text-align: center;">It's not much of a problem.</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">I'd second that.</div> </div> <div style="display: flex; justify-content: space-around; margin-top: 5px;"> <span>621</span> <span>622</span> </div>
B	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; padding: 2px; text-align: center;">It's not much of a problem.</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">I'd second that.</div> </div> <div style="display: flex; justify-content: space-around; margin-top: 5px;"> <span>621-1</span> <span>621-2</span> <span>622</span> </div>
C	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border: 1px solid black; padding: 2px; text-align: center;">621-1</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">621-2</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">621-3</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">621-4</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">622-1</div> <div style="border: 1px solid black; padding: 2px; text-align: center;">622-2</div> </div>

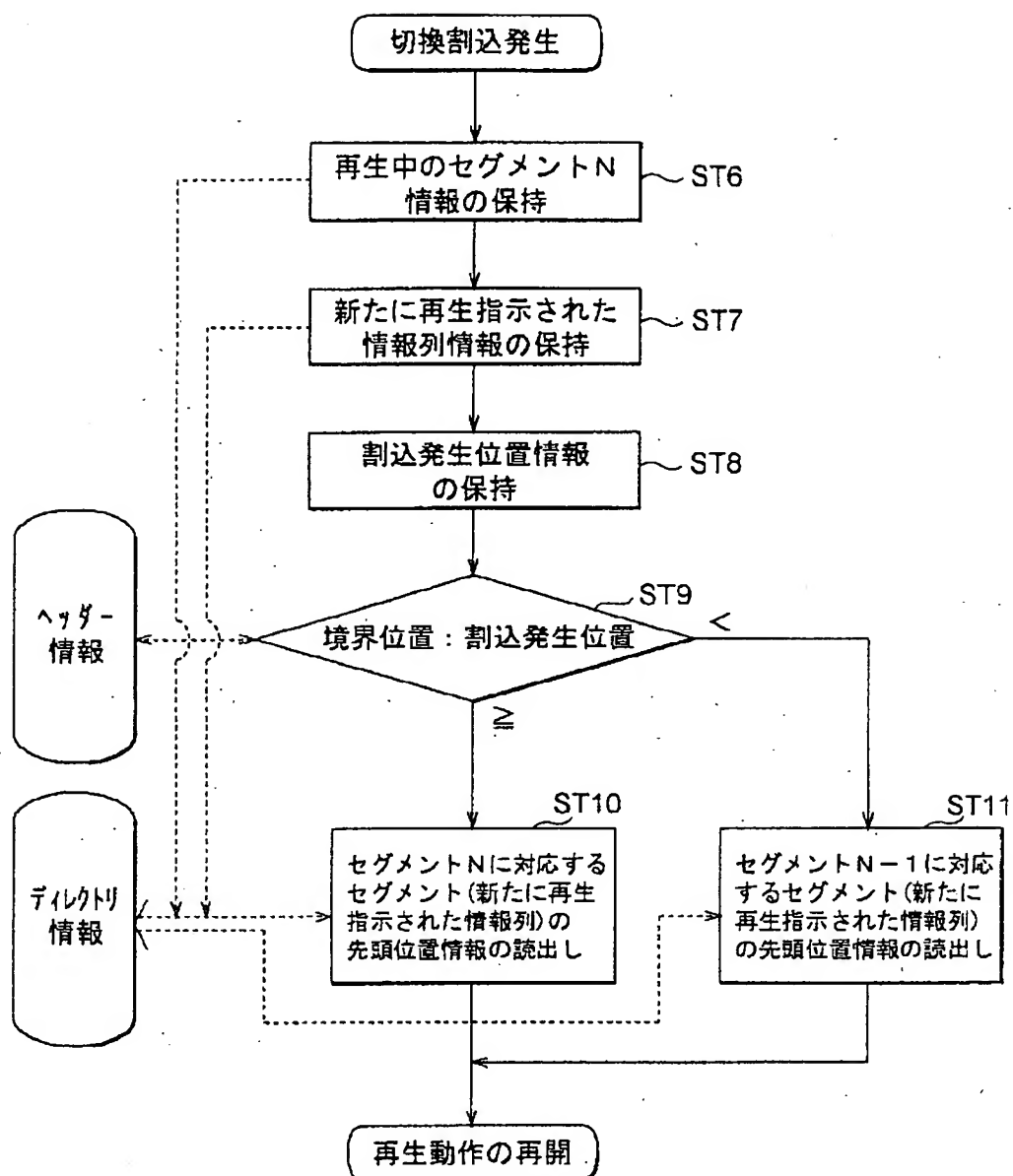
[Drawing 16]



[Drawing 17]



[Drawing 18]



[Translation done.]